

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

Sent: Tuesday, July 05, 2016 3:39 PM

To: pds4ddwg

Subject: Notes from PDS DDWG 2016 06 30

title: DDWG Notes 2016-06-30

layout: default

date: 2016-06-30

June 30, 2016

Notes by Debra Kazden

Known Attendees: ****EDIT STILL** *******

R.Alanis, M. Gordon, E. Guinness, S.Hardman, L. Huber, S. Hughes, C. Isbell, R. Joyner, D. Kazden,
L.Neakrase, J. Padams, A. Raugh, B. Semenov and J. Stone

Meeting Agenda and Summary

1) SCRs Under Discussion:

-- 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
- 20160617: updates discussed by DDWG
- 20160621: Jordan waiting for feedback / comments; then will go back to DDWG
- ** (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-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; E.Shaya led discussion; 2 competing implementations
 - address 2 issues:
 - bang for buck in terms of worth doing
 - implementation recommendation
 - WG: Steve, Anne, Jordan & Chris, Lev
- **(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-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
- 20160512: emailed Steve asking if the WG had met and is SCR ready to go back to DDWG for discussion?
- 20160615: emailed Steve asking if the WG had met to resolve issue
 - ** (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
 - 20160623: Under DDWG review
 - ** (Discussed) **
- CCB-159: Bug fixes for Version 1.7.0.0. (J.Hughes)
 - 20160426: Open & Under DDWG review
 - ** (Not Discussed) **
- CCB-162: Move <md5 checksum> from Object Statistics to Array (R.Simpson)
 - 20160622: Open & Under DDWG review
 - emailed Steve to TA
 - 20160629: TA'd; back to DDWG for discussion
 - ** (Discussed) **
- CCB-163: Correct the Instrument.Type enumerated list (L.Huber)
 - 20160622: Open & Under DDWG review
 - emailed Steve to TA
 - 20160629: TA'd; back to DDWG for discussion
 - ** (Discussed. Voted. Will be sent to CCB.) **

Notice sent before the telecon in email from R. Joyner - June 22, 2016 See enclosures:

-- a list of the full topics under discussion by the DDWG

CCB/SCR Statuses:

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

-- 20160616: Ready; sent to CCB (as requires IM change)

-- 20160628: Queued_for_Implementation

-- Consent item PASSED: 5 YES (ATM, GEO, IPDA, PPI, RMS); 2 No comment (IMG, SBN)

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

(Please review and be prepared to discuss and possibly vote:

-- CCB-163: Correct the Instrument.Type enumerated list (L.Huber)

-- 20160622: Open & Under DDWG review

-- 20160629: TA'd; back to DDWG for discussion / vote

-- CCB-162: Move <md5 checksum> from Object Statistics to Array (R.Simpson)

-- 20160622: Open & Under DDWG review

-- 20160629: TA'd; back to DDWG for discussion / vote

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

-- 20160418: Open

-- 20160623: Under DDWG review

SETI Issues (M.Gordon, et al)

-- Reconciling Local Data Dictionary Versions with New System Builds:

-- see enclosure for real-world example of (SBN-PSI LDEX) xml label:

-- Consistency with the common dictionary

-- Schema validation provenance

-- Version identifiers

-- Configuration management

-- Use of XML Catalog files in archived bundles Exposing elements in the PDS4 Common XSD

(S.Hughes)

-- Discussion on elements to be exposed

-- review emails from Steve: Exposing elements in the PDS4 Common XSD

-- Scheduled to be implemented in development version of IM; possibly v1710 ?

DDWG Telecon

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

This went to the CCB. Passed. Queued for implementation.

~

Hope everyone has reviewed CCB-162 and 163. CCB-163 was updated - all updates have been TAed.

~

CCB-163 - Correct the Instrument.Type enumerated list See <https://pds-jira.jpl.nasa.gov/browse/CCB-163>

Question: Any questions on the SCR as written?

Answer: Crickets.

We will be voting on the SCR as written - to endorse or not so much.

****The Vote - to endorse the SCR as written** **ATMOS -**** (Comment - Even though I am the author,

my vote is to not endorse. Have two issues with the SCR - 1) Dick's item three in jira comments (which reads: "I am concerned about the granularity differences among the types — especially Atmospheric Sciences, SBN Exceptional Instruments, and Regolith Properties. The definition of the last encompasses all instruments of interest to geosciences, goes well beyond regolith instruments, and could subsume half of the other types in the list.") should be pretty obvious - that things in the enumerated list do not have as much specificity as would like - but hopefully, we can live with it.

2) Search is dead. No possible way to do search - except maybe for fields and particles.) **Not Endorse EN - Endorse IMG - Abstain GEO -** (Comment - Understands ATMOS's issues, but would rather endorse and get this out of the way then to spend any more time on it) **- Endorse PPI - Abstain NAIF - Abstain SBN -** (Comment - Not feeling this is a good solution - was going to abstain, but that feels washy-washy, so voting no, we are not ready to roll here) **- Not Endorse RS - Not Present Rings - Endorse IPDA - Not Here**

This absolutely leaves us in a quandary. This is really not good. We could withdraw this and never discuss it again.

~ That's not an option. Not with this list.

Three votes to endorse.

We can send this to the CCB with this vote and let them work it out.

(Note taker missed a few details of the conversation because she was being addressed. Someone suggested that PPIs vote be changed.)

Question: Is Steve Joy lurking? Could Debbie get someone at PPI to vote? An email by the end of the day?

Answer: Debbie will talk to Todd.

Maybe people who abstained should write a written explanation. Abstaining is abdicating responsibility.

ATMOS should write and list why they feel this is not useful. Include some use cases.

~ ATMOS has given several use cases in the past to show why this won't work. Not again.

~ Someone wants to look at use cases and try to understand.

Part of the reservation is waiting for a presentation from EN on how the registry actually does work and how that would affect this. Without that, have to assume that the superficial details underlie it all. I don't believe it, but can't assume this solution will address the larger problems.

Question: Last time, we agreed that a hybrid solution is the only possible solution. Do we still want a hybrid list with two levels of specificity?

Answer: Someone can live with it, but doesn't think it's the right solution.

~ Then someone else come up with one.

Another Question: Are we saying this is no better than anything we had in the past? Every time we present something it's not good enough. Think this was better than we had in the past. What's the goal?

Answer: Something useful for search.

~ Then we are back to use cases. Want to address each one. The CCB will ask the same questions. Would like to work with the use cases and see if being met.

So, here's where I want to go. We will wait to see if PPI changes their vote. Either way, want to send this to the CCB to hash out. Will tell them it was hardly endorsed by the DDWG. Working on use cases is fine.

****Action Item - Debbie**** check on the PPI vote.

Question: Does anyone have any last words?

Answer: (Silence)

CCB-162 - Move <md5 checksum> from Object Statistics to Array See <https://pds-jira.jpl.nasa.gov/browse/CCB-162>

RS has approved the SCR as written.

Question: Anyone have any comments?

Answer and Another Question: Added a late response to jira. Would like a clarification the outcome of this passing. Would that result in being able to store checksums for file or image but not other objects? Is that right?

~ So... it adds MD5 checksum as an optional attribute in array...reading the SCR... would have to go back.

Another Question: Object statistics could be about a binary table or other things, but 10.25 in the IM is only arrays. So, can a table have object statistics?

Answered with Another Question: Thought that MD5 checksums was moving up to array. Is it not?

Another Question: Can you have it for other objects?

Answer: It should be in the parent class of all digital objects.

~ Thought the proposal was to put it in base object - array. Then it would be inherited. Not proposing it be put in table.

There is a work group working on this. They have been exchanging email. Not ready for a vote. Not Updated.

~ Someone thought this was ready. Work group should work this out and address new comments in jira.

~ Fine.

Question: The question is if we want to be able to store checksums for objects besides array.

Answer: Wouldn't want a field in a table - would want at the highest level.

~ Byte stream might be too abstract - maybe.

~ Looking at the IM - wouldn't put it in byte stream - would put it in array or table base- in complete files. File checksums rather than digital object checksums.

~ Someone agrees.

~ It could be optional. Those are just two bases where it might be desired to have a checksum.

Question: Could people email their comments on this to the work group for their next meeting?

Another Question: Reply in email, not in jira?

Answer: Someone didn't know there was a work group. Would prefer email - not clutter up jira.

Question: Who is in the work group?

Answer: Ron, Dick, Anne, Chris, Steve and Jordan. Anyone else can join. Will send them the email conversation.

Question: Anything else on CCB-162?

Answer: (Silence)

CCB-156 - Inconsistent Discipline Dictionary Technique for Local Internal Reference, et al.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-156>

Prefer not to go discuss this and go on to SETI.

CCB-156 is a major issue from SETI.

~ Not major, it has relevance to SETI.

Question: Test case to expose local internal reference as a type?

Answer: Yes, but doesn't solve the entire problem. Still user and DP confusion on where to look for it.

There's a policy question on how LDDs reference classes from core namespace. Two issues. Uniformity and non-backwards compatible change.

Will send schemas with exposed solution for people to look at.

~ If LDD tool is the portal through which all changes are made - will have issues to deal with. Need to make sure everything is in place.

Question: When we get the new schema, will there be modified LDD tool that makes correct

namespaces? We do that now.

Answer: Problem - LDD tool prescribes local internal reference to common.

~ Yes - it will include updates.

Question: Are we done with CCB-156? Have a short term plan?

Answer: Not looking for an instant solution to that.

SETI Issues

Someone's view of the world, before last week, thought it was all theoretical. Reviews a lot of bundles. Sees an adherence to EN's best practices. All the bundles have fallen under the umbrella of what EN wants, so wondered what's the big deal about SETI. Everyone was doing what EN wants. Then here comes a bundle - from SBN-PSI, the LDEX bundle - and now the real world is not following EN's best practices. So, we have all these bundles, and they all look the same - and then an outlier. Disappointed. It was suggested it might be made to look like the others, but that comment was ignored.

Someone hasn't seen the LDEX bundle. Doesn't know the history or how it was built.

~ Will send the link to the bundle to the DDWG.

Question: Are most problems in the sample label?

Answer: Sending the URL now. Been watching bundles as they come in to EN. Mostly happy, and then this arrived. Let's give people an opportunity to review this.

Question: Are the concerns with the bundle structure, labels..?

Answer: Labels principally.

~ Looking... The identification area - the title is the name of the file, IM is 1.1.0.1 - very old - everything is upper case. My guess would be that this is output from a tool that converts PDS3 OLAF tool data to PDS4. Observing system is weird. Things look incorrect - maybe because of the old IM.

NAIF will be using and old one forever. (Speaker was very difficult to hear) NAIF wants feedback on if that's a problem. The way things are changing, there will always be non-backwards compatible changes.

Things are being deprecated. There were structural changes from 1.1 to 1.3. Where there are structural issues will have to migrate things that are earlier than 1.3. The title is relevant because it's returned where users search for data. Not always useful if it's just a file name. NAIF may design so that the title for an individual kernel might be enough, but in general thinks title should have meaningful human readable information.

Lets' take the NAIF issue off line.

SBN-PSI has no objection to updating the LDEX bundle. They have made changes that could fix things.

LADDEE was a very early mission so not all of the best practices were known yet.

SETI Issues - one that someone would like to vote on is the use of XML catalogs in an archive bundle. Last week we had general agreement that we would never include them in an archive bundle. Last week we were hard pressed to find something to discuss. Would like to vote to agree or not agree that XML catalog files - that use of XML catalog files in archive bundles - yes or never.

Question: Why? Why not?

Answer: Been hashing this out for weeks.

~ Catalog file merely provides a local file location for file in a label. PDS4 schema in a local directory.

Argument is that shouldn't be the case. All bundle references should use URI namespace information.

Don't want catalog file in archive bundles.

~ Got it. In a time warp. Sorry.

Simple. Question is should we prohibit catalog files in an archive bundle. Yes means not allowed, no means allow them.

**The Vote - Yes or no - allow XML catalog files? Yes is allow catalog files and no is don't allow.

Rings - No

RS - Not Present

SBN - No

NAIF - Abstain

PPI - No

GEO - No

IMG - No

EN - No

ATMOS - No**

We have agreement - a step in the right direction. The question is how to impose this.

~ Peer review or maybe a tool.

~ Need an SCR to express this rule in the SR.

~ Okay.

~ Can't write it in schematron.

~ Right.

Question: Couldn't we name it a specific way?

Answer: Need to open the file to read it.

Another Question: Don't you need "xml:catalog" to activate it?

Answer: No. We already have something that flags if you try to validate. More to the point, can't actively prevent people from including it. If someone wants to label it as a document, it's not worth our time...

~ Someone agrees. It's a best practice. Maybe a peer review best practice.

~ Referential integrity has to be there without the XML catalog file.

~ XML catalog file doesn't get you referential integrity.

The SCR someone would like is that the referential integrity of a bundle can not depend on a catalog file.

- ~ Have to be able to find PDS4 common schema without a catalog file.
- ~ The SR should never mention catalog files - ever. Validators shouldn't use them for ingest validation. Central registry should know how to use canonical schema. Want to start with namespace.
- ~ Someone wants to get rid of it and say the referential integrity of XML files should not depend on catalog files.
- ~ EN validation doesn't use catalog files.
- ~ But it can. We can't tell people how to validate. Referential integrity has nothing to do with catalog files. Namespace references are the model.
- ~ We are trying not to confront... Validation has to be achievable without XML catalog files.
- ~ That's not a statement we can make.

We voted, so the question is where the statement goes.

- ~ Maybe the DPH. No XML catalog files in the archive. Can tell DPs we won't ingest them.
- ~ You can tell people where to be at 10:00am, but you can't tell them how to go - by bus, train - only can tell them the goal. Not how to get there. Eclipse doesn't use XML catalog files. Can't tell people how to run their validation, can only tell them what the result has to be.

Question: Is this a requirement or a guideline - no catalog file?

Answer: Someone wants it to be a requirement.

Another Question: So, a standard we can enforce or a guideline? If it can't go in the SR - where does it go?

Answer: Sounds like the DPH.

- ~ Someone had hoped to get to the ramifications of not allowing catalog files - so we could have best practices - failing.
- ~ We got an agreement. Just need to decide where we get it.

The intention wasn't to have an SCR that says thou shall not use catalog files ever.

- ~ That should be include instead of use.
- ~ That's what we all agreed on.

Question: What's the next step?

Answer: It's 10:30 am. Ron and Steve have to go. We didn't get very far.

~ We're at the end of the agenda.

~ We didn't get to exposing elements.

Back to LDEX. There were a ton of review comments. Looks like the comments were ignored.

~ No, it looks like it was updated.

~ PSI person thinks they did update based on the comments.

~ Maybe looking at old data.

~ PSI added some data. The original was 2013.

Question to PSI: Can you send a link to the most recent? I clicked download and what I got was not good.

Answer: It's the most recent, but still not good.

~ Thought LDEX was frozen at 1.1. So maybe that needs to be discussed.

~ Need to see the review comments.

Question: So, should this be sent to NSSDC for testing?

Answer: Would think a version 1.1 label would be problematic for NSSDC.

~ LADEE is all 1.1

~ Should check with Carol Neese.

~ We want to test NSSDC.

~ Carol is willing to migrate to a later version - it's just not done yet.

Next week.

No ATMOS people available next week.

~ We will meet anyway.

Send burning SETI issues to Ron.