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

Sent: Tuesday, May 17, 2016 10:37 AM

To: pds4ddwg

Subject:Notes from PDS DDWG 2016 05 12

---

title: DDWG Notes 2016-05-12

layout: default

date: 2016-05-12

May 12, 2016

Notes by Debra Kazden

Known Attendees:

R. Chen, M. Gordon, E. Guinness, S. Hardman, L. Huber, S. Hughes, C. Isbell, R. Joyner, D. Kazden, T. King,

J. Mafi, L. Nagdimunov, L. Neakrase, J. Padams, A. Raugh, R. Simpson and J. Stone

Meeting Agenda and Summary

1) CCB/SCR Statuses

The results of the 2016-0510 e-vote are as follows:

-- CCB-132 Units of Map Scale is Badly Defined

   E-vote PASSED: 5 Yes (ATM, GEO, IPDA, PPI, RMS), 2 Failed to vote (IMG, SBN)

-- CCB-143 Validate field format via regex

   E-vote PASSED: 4 Yes (ATM, GEO, IPDA, RMS), 3 Failed to vote (IMG, PPI, SBN)

-- CCB-158 Restore J2C as valid format for supplemental data

   E-vote PASSED: 5 Yes (ATM, GEO, IPDA, PPI, RMS), 2 Failed to vote (IMG, SBN)

    (Brief Discussion)

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 ?

   -- 20160512: DDWG discussion topic; TA'd and up for vote

   (Discussed - being sent to CCB)

-- 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-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

(Not Discussed)

-- 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

-- 20160511: 2nd email to Steve to ask for TA; to make ready for vote

(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 ?

   (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-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

   (Discussed)

-- SETI Issues (R.Simpson et al)

   -- Status

   -- Issues in XLS have been vetted by SETI notes

   -- Issues to be "consolidated" & prioritized

   (Not 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 11, 2016 See enclosures:

1. a list of the full topics under discussion by the DDWG

The results of the CCB 2016-0510 e-vote are as follows:

   -- CCB-132 Units of Map Scale is Badly Defined

        E-vote PASSED: 5 Yes (ATM, GEO, IPDA, PPI, RMS), 2 Failed to vote (IMG, SBN)

   -- CCB-143 Validate field format via regex

        E-vote PASSED: 4 Yes (ATM, GEO, IPDA, RMS), 3 Failed to vote (IMG, PPI, SBN)

   -- CCB-158 Restore J2C as valid format for supplemental data

        E-vote PASSED: 5 Yes (ATM, GEO, IPDA, PPI, RMS), 2 Failed to vote (IMG, SBN)

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

(1) Please review CCB-65 including the TA and be prepared for a vote:

-- 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 ?

      -- 20160512: DDWG discussion topic; TA'd and up for vote

(2) In reviewing CCB-152, the SCR as written doesn't address adding the POSIX to the IM

      -- shouldn't adding the POSIX be included in the set of requested changes ?

      -- if not, then the SCR (as written) has been TA'd and is ready for a vote ?

-- 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

-- 20160512: DDWG discussion topic; TA'd and up for vote

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

    -- Status & develop implementation plan

-- Exposing elements in the PDS4 Common XSD (S.Hughes)

    -- Discussion & implementation plan

DDWG Telecon

CCB/SCR Statuses

The first thing at the top of the agenda is that three SCRs went to the CCB. CCB-132, CCB-143 and CCB-158 all passed.

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

On to the one from Anne that we want to vote on.

Question: Does anyone want to discuss this?

Answer: Sounds fine to someone.

~ Someone else keeps thinking this is done.

We will close the discussion on this. It is going to the CCB - endorsed - going in two weeks.

Question: Are they on a two week schedule?

Answer: No, but their schedule sort of conflicts with ours.

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

This is about the POSIX and field format stuff we discussed last week. The SCR as written doesn't address adding POSIX to the IM. There's been back and forth on if we should do it as is. We're voting on that and go back to it later.

~ Someone doesn't understand why that's not automatic.

Surprised when the TA was being done that it wasn't there - thought we should discuss it. Didn't want to make an assumption.

~ In the future, should have the option of asking the DDWG for expertise or suggestions on implementation.

~ Person doing TA has no authority - if they see something missing they send it back to the triage group and then it goes back to the DDWG. Person doing the TA just looks at the requested changes to see if it makes sense.

~ Maybe this needs to go back to the MC for a procedural discussion. Seems there's a step missing.

~ Someone thinks the process is working. We caught this and asked the author about it. Thinks it's right for for them to be TAed as written. (For this one) we can vote to go as written or the author can edit it.

Someone keeps bumping into the idea that the person who submits the SCR is also responsible for submitting the solution. If we need something more than the CCB can ask the DDWG for more, but that's a waste of time. There's a big discussion in the MC now about the lack of efficiency. The issue here is what's missing from this SCR. Looking at the SCR, doesn't see the problem.

~ The suggested changes don't ask that POSIX regex be added to the IM, so in the final review for the TA, saw it was missing and sent it back. Surprised at the idea that the TA person could solve this.

~ It shouldn't go all the way back to the author.

~ Yes it should. Not everything has to go to the DDWG. Should have asked the author if he wanted to

add it.

~ That's what we did.

~ Someone disagrees. People get targeted when things go wrong. TA person saw it as a red flag that needed more review.

~ That's not what was said.

~ We haven't voted on this yet - seems it could still be worked and improved.

~ All someone does is write the TA. Has no authority.

~ That leaves it to the DDWG.

Question: So it can only be discussed by the people on this telecon?

Answer: No. DDWG just makes an assessment.

Another Question: Is this going to the CCB to decide if we should design a solution or to implement a designed solution? Seems there's a missing step - would think the TA was the second step, not the last step. Doesn't seem to be a design and test step...?

Answer: There is a design and test step. When write the TA that's done, but in the last few months been TAing before the DDWG look at things - then the TA is modified - it's the last step.

~ So, the TA is a work order. No judgement on merit or seeing if things were left out. So, second step is to TA - CCB gets to access if it's worth the order.

~ Yes. TA is written based on the discussion so far - then it goes to DDWG - then work order is sent to CCB.

The problem is the yada yada yada between the steps. Things are sent back to the SCR writer, not to the DDWG.

Someone asked the author of this SCR if they wanted this voted on as written or if they wanted to edit it. They're the author.

~ That's not good configuration management.

Someone does not want to have this discussion.

~ Then someone else is going to the MC.

Someone likes the idea that for complicated ones we have work groups. Agrees that the person who identifies a problem is not responsible for coming up with the solution.

We're debating the process - not this SCR.

~ That guy is right.

Question: Is this ready to go to the CCB?

Answer: The process is not working so great. For this SCR, decided to leave it out, but when someone else pointed it out again then thought that it did need to have more discussion. Doesn't think it's ready. This group should finish it - fix it - then vote. The issue, the reason he left out the specification - formation rule - was that it implies that there are no other restrictions.

Looking back, thinks it should point to the SR - say subset is allowed - can't put all of the SR text in the IM. Thinks that's the solution.

Question: Any objections?

Answer: People agree.

~ The text will be rewritten for next time **(Action Item - Lev)**

There are still another two issues. One is field format is optional everywhere - need to decide if we want that required. Also, there's the issue with right justified - float point issue too.

Question: Do we need a separate SCR?

Answer: It shouldn't all be bundled into CCB-152. CCB-143 is already complementary.

~ The SR description is fine, the issue is in the IM.

Question: If you say that field format of F.5, must there be five digits in the field?

Answer: We need to discuss this. (Note taker missed something - value 123.456 - padding) SCR leaves it

a bit ambiguous. Can Validate without requiring it. Would like comments.

~ Someone absolutely wants to require padding after the decimal - prefers leading blanks. Trailing zeros is a requirement.

~ That requires more precision then might be there. It varies across data.

~ That happens a lot at SBN. Don't want to over represent precision of the data.

~ Didn't realize there is data with mixed precision in a single column.

~ Not a lot, but it does happen.

So maybe we want to leave the text as is. Everyone should read it.

Question: The SR, for E, specifies you can have two numbers - do we want to change that? Maybe for three numbers?

Answer: It has to do with if the number is really small or really large. Univax has some very big numbers. We can probably live with two digits and deal with future cases if they come up.

~ We may want to consider if the units are inappropriate.

~ Also need to consider backwards compatibility.

So, not making any change to the SCR except to specify that you should look to the SR.

~ Someone also has a few grammatical suggestions.

Don't think we're ready for a vote.

Proposal: CCB-1xx: Remove Enumerated List from Instrument.type See email from Steve Hughes, May 11, 2016, Report on the Instrument Classification Exercise

Steve sent out a report. The goal was to determine instrument types and find a less intuitive way to control the list. The mechanism did work. The results are a move toward illustrating the potential. There's clean up required - didn't do any. The point here is - eighty percent of the work is clean up, but we have a possible path forward. There are two lists - they show we can control specificity.

There's the short common name and you add value to get the common name. For example, add wavelength to get visible imager rather than just imager.

Question: Feedback?

Answer: Someone has a lot to say, but will let others go first.

~ Someone else is not sure which list to comment on, but seems still a mix of specificity in both of these lists. Not sure what the actual values would end up being. Short names may not find everything in search.

~ This wasn't done to implement search. The process is important here, not the results. Still need a lot of clean up. There's a wide variety of things from PDS3 - discriminators will help drop things out - will make specificity more similar. All those issues would go to the next level. Names don't matter here. The second time around would make sure we use the proper terms.

~ If we use this the definitions of the discriminators needs to be cleaned up. People read them different.

~ Yes - we used what was initially proposed, but we do need to firm up the definitions. For this exercise absolute correctness wasn't the issue. It was to make sure of the process. Need to make sure multiple nodes classify things the same way - have agreement on what's being measured. Need to clean up the discriminators.

Regarding the 170 that weren't classified - some will fit in. There's more work. The 170 will be hard, not impossible.

~ This is modeling, not implementation, so not worried about search.

For PPI, whatever classification method you use - for PPI there are only five answers.

~ Took the answers - radio science and magnetometers etc., those are verbatim in common name - the question is the short name - were easily mapped into the process. PPI was the simplest, SBN had a lot more. They have a lot to resolve. ATMOS too. Have to decide if they will break the packages apart - things that measure atmospheres, things that measure temperature, etc.

It seems this answered a different question than the one asked. The question was why this has to be an

enumerated list. Seems you took an enumerated list and reduced it - don't see that it improved the situation. Might have gotten specificity to equivalent level, but still no complete list. Doesn't address what makes good enumerated list. Seems like this creates serial numbers- we already have that with LIDs. Wants to go back to the three possibilities (See slide 10 - Way Forward - EnumeratedListsB.PPTX attached to meeting agenda for Apr 21, 2016- sent April 19,2016, by Ron Joyner)

"1 -Create a multi-tiered instrument.type with a small enumerated list at the top level which has matching specificity levels and a schematron controlled lower level to provide flexibility and ease incompleteness worries.

2 -Set aside the current enumerated list and start over requiring the nodes to provide the components that would serve their community.

3 - Make the list unenumerated. Trust the nodes to instruct providers what proper values should be."

Think multi-tiered should have half a dozen things at the top - then under categories you have unenumerated with a steward who controls the values that get added.

~ Thought the exercise was coming up with the top level. Don't think it will reduce to half a dozen, but it will be more broad than imagers versus spectrometers.

~ Imagers/spectrometers would be a category along with atmospheres, geosciences, fields and particles, then go on from there - maybe other.

~ Someone thinks the process is choosing the discriminators. Should never use other.

~ Agree, but sometimes it's the only reasonable solution.

~ Implementation is ignored - this is a modeling problem. Trying to make a list of instrument types. No control in PDS3 - never ending growth. This will not expand without good reason. No longer intuitive, but we would know why we need to add something. Also addresses specificity. Thinks on the same page, but unsure if the PDS3 list includes all possible instruments. We can add new ones, but...

~ We already know two new instruments without PDS3 comparables.

~ We should use those as test cases.

Question: Other comments?

Answer: Thermal probe and Seismometer.

~ Mars 2020 will have additions too. All new missions will probably have new things to add.

Someone says this is just modeling, but it is about implementation.

~ This is a list - the issue is we have short text fields - could leave that alone. We have Primary Results Summary - won't solve problems without enumerated list. If the issue is if things are facets or something else - that's up to Sean. We shouldn't use constraints on implementation when figuring out the hierarchy. The discriminators - PPI had a perfect example for Primary Results Summary.

~ That's nice for PPI with their four instruments, but others have more.

~ The size of the list doesn't matter.

~ Size does matter.

~ Can control size with the discriminators.

~ If everyone saw there's a new instrument and we all used the same algorithm to travel through the tree, should all end up with the same instrument type.

Question to ATMOS: This started with your issue. Is this getting resolved in the time frame you need?

Answer: If this is not fixed in time for the next build then we have a problem. Something needs to go to the CCB no later then the end of August.

~ We need the solution by then. Don't want them to kick it back here.

~ Need a solution we can all live with.

~ PPI doesn't have a dog in this fight. If others work out a solution that allows the solution we want then PPI doesn't need to be engaged in all the details of thinking about how we move forward.

~ Someone understands ATMOS's desire for a short list. Has a few basic questions to answer before we move forward. Not sure what a short list looks like. Also need to know if we are breaking apart packages.

~ We should very rarely ever have to add new values. Yes, this classification scheme addresses that. Just not sure why short versus long.

Every time a new mission is selected something will be added.

~ Let's make testing that the next step in this exercise experiment.

~ We need to go around the table to see what people think before more man hours.

~ Someone is interested in how much time people spent on this. Took EN about 16 hours, but now maybe updates will maybe be one hour per node - so maybe eight hours. Interested in how long the nodes took.

~ Atoms - Took two hours - didn't do it very well. Maybe four hours to get it exactly right.

~ GEO - Spent about eight hours - not sure all correct because of the discriminators.

~ IMG/USGS - eight to ten hours, but not sure that it would be the same.

~ Might take more time now.

Where this all comes from - where we're going - SCR was that we need an unenumerated list - CCB agreed not tenable, but sent to DDWG to see if there's a better way. We can't answer that yet. Thinks we need to continue.

Question: What about the multi-tiered list? What if we put this effort aside and try that method?

Answer: Valid.

~ Someone thinks we're talking about two multi-tiered lists. Thinks this is a multi-tiered list. Would like a draft. Confused.

~ Yes (draft will be sent) **(Action Item - Lyle)** Seems like this could have an impact on migration - on context products.

Again, someone would like to hear from the nodes - if we should continue to let Steve move forward, try the other method or push for the CCB to make unenumerated. Agrees we're not really there yet, but that's his opinion.

~ Someone would like to have Steve continue. Would also like to see Lyle's tiered list. Not sure, seems two approaches to the same solution. Thinks opening the second frontier is a good idea.

~ Someone else is okay with letting Steve run on.

~ Another person is still at a stage of thinking analytically - not clear in their mind yet. Thinks will have abstract conversations with Steve.

~ Another someone else is not sure what Steve's next step is - and not sure what Lyle's list will look like. Not sure having node instruments is a good idea. We need to be able to get new instruments included in

a timely manner. Don't want to have to wait for the new IM version. People got burned in PDS3.

~ Yes, people got burned in PDS3. This new mechanism would get things in quickly. We need a quick way to make changes - but it's not the same problem.

Someone wants to go back to the purpose of this attribute - if we're sorting into bins, and if for search, than values need to be useful for search.

~ Different issue.

Question to GEO: To clarify, are you not wanting enumerated list?

Answer: At this point yes - haven't seen a viable alternative.

~ IMG is on the fence. Don't want missions calling the same instrument different things.

~ Someone else is also on the fence, but sees the value added - cost to solve this doesn't seem worth the value. Not sure it's worth the man hours.

~ Yes, cause it just ends up in the context products - not being used in the system.

~ Yes.

~ Yes, it is being harvested and used in the query model which is implemented in version 1.6 Query model is not yet completely implemented in product observational.

~ Observing system component type has nine possible values - instrument, laboratory, etc. Things like camera are wrong.

~ Instrument type is required.

~ In the context product.

~ Also harvested and used in search.

(Rings hung up)

IMG is willing to go forward, but not to much time. Also okay with removing enumerated list. Also needs to sign out.

~ Todd also has to go.

Let's give Steve one more week - next week we can decide which way to move forward.

~ Dick and Anne won't be here next week.

~ Fine with GEO.

~ Some clean up can be done in a week, but would need node input, so two weeks to see if the classifications will solve the problems.

Okay - Ron's meeting.

~ (Silence)

Question: Did he get called away? Steve should take over.

Answer: We were going to go on to exposing classes...

~ Someone is not sure he understands and not sure we should precede - seems we just lost a bunch of people.

Exposing elements in the PDS4 Common XSD

We can answer questions. In the common model we define things as type or element. In Oxygen element gives a pull down menu. Element is the real thing. From version 1 to 3 we exposed all their types - made them elements - went from one schema to reference to something in another. Before 1.3, just exposed the element. After 1.3 we changed things. Elizabeth screamed that things weren't working - there was an issue with three or four that should be exposed in common, so people can reference them.

See email from Steve Hughes, May 11, 2016, Exposing Selected Classes

(Note taker missed some)

If you want to reference internal reference or ? or ? can't - just can get the type. Closer to what ATMOS wanted years ago. Ability to use a class in an LDD.

Question: Simpler? And in new current release - impossible?

Answer: Good point. That's another issue. Yes. Not an XML schema expert, some say ways around it, but there is a real issue here.

Question: Other opinions?

Answer: (Silence)

Comment: Wow. That was a rousing discussion.

We can discuss this next time.

~ Maybe it would be helpful if we heard the downside too. Unsure if there is a problem.

~ In tools like Oxygen, pull down list would be longer. There's been email back and forth about this.

~ Pull down isn't a good argument, but that's the way schema works. Only elements can be cited in others type definitions.

Confusion.

Potential downside, but could require all LDDs to be written with LDD tool and avoid abuse. LDD tool's programmed to provide correct name/type. Having element names defined changes how it looks - also enforces that in all LDDs with PDS internal reference have the same requirements/restraints everywhere. Can be done in less annoying ways than expanding the list in Oxygen.

We can repeat this next time.

~ Yes. We should discuss this.

Meeting called to a close.