

title: DDWG Notes 2017-01-05

layout: default

date: 2017-01-05

January 5, 2017

Notes by Debra Kazden

Known Attendees:

M. Gordon, E. Guinness, S. Hardman, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, L. Nagdimunov, L. Neckrase, J. Padams, A. Raugh, R. Simpson and J. Stone

(The full meeting agenda attachment that usually accompanies the meeting announcement was not included this week.)

Notice sent before the telecon in email from R. Joyner - January 3, 2017

****CCB/SCR Statuses**:**

-- None

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

(1) Please review and be prepared to discuss:

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

-- 20150922: Open; under DDWG discussion

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

-- 20160929: back to DDWG to form WG; candidate for sunset clause

-- 20161010: from Anne: Jess Stone (jstone@psi.edu) will be the SBN rep for this

-- need participants to form WG

-- 20161013: WG formed: J.Stone (Chair); Lev & me (participants)

sent email to Dick & Jordan to ask if they will participate on WG

-- 20161026: sent email to Dick to see if discussions between Dick and J.Stone are solid enough to update SCR

-- 20161026: Dick to update SCR; will split off NaN & INF; to be addressed in separate SCR - see CCB-170

-- 20161109: Dick has made extensive updates to SCR; needs Steve to TA

-- 20161109: At DDWG, Dick dropped out of WG; Lev/Anne to define competing implementation(s); send to DDWG

-- 20161129: In wait-state until can attain consensus within SBN

**** (Postponed until Anne is available) ****

-- 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

- 20160929: back to DDWG to form WG; candidate for sunset clause
 - on hold; not high priority; too few people resources
 - bring back to DDWG: Dec 2016
- 20170103: back to DDWG discussion;
 - ** (Postponed until Anne is available) **

- CCB-168: ASCII NonNegative Integer may or may not include + sign? (L.Nagdimunov)
 - 20170103: Open & Under DDWG review
 - ** (Discussed - decided it is a bug fix) **

- CCB-170: Deprecate IEEE 754 NaN and Inf in Favor of PDS4 Special Constants (R.Simpson)
 - 20161110: Open & Under DDWG review
 - 20161129: in wait-state until CCB-131 is resolved
 - ** (Very Brief Discussion - waiting on CCB-131) **

- CCB-171: Split hardware-compatible ASCII numeric types from the unbounded (L.Nagdimunov)
 - 20161012: Requested Lev provide cogent statement of the issue to be resolved; send to DDWG to start email exchange & discussion
 - lots of email discussion; inconclusive as to whether issue is converging or diverging ?
 - 20161026: Can a problem statement be written into an SCR where there is some measure of certainty that this is the issue to be resolved ?
 - at least one email questioning whether this issue is worth the time / effort to discuss / resolve
 - 20161109: Lev requested another week to work issue; before additional DDWG discussion
 - 20170103: Open & Under DDWG review
 - ** (Discussed - will be TAed so we can vote next time) **

(2) Email exchanges and other stuff

- PDS4 equivalent for the PDS3 keyword SOURCE PRODUCT ID (S.Slavney / J.Padams)
 - 20161129: Open; under DDWG discussion; form WG
 - ** (Discussed - IMG and SBN will send examples to DDWG) **

(3) Product Update (M.Gordon)

- 20161109: Mitch sent out Product Update email; send to DDWG to start email exchange & discussion
 - 20161129: Mitch to send out example(s)
 - ** (Not Discussed) **

DDWG Telecon

Steve sent an addition to the agenda - Composite Structure. (See emails from J. Hughes, Jan 4 and 5, 2017)

Nothing since last meeting. No SCRs went to the CCB.

CCB-138 - Mismatch between context object types and values of <type> in <Observing System Component> class See <https://pds-jira.jpl.nasa.gov/browse/CCB-138>

We had left this on the back burner.

Question: Do we want to discuss this now or wait for Anne?

Answer: Anne is on vacation. Need to look at it.

Postponed until Anne is back from vacation. This is one of her favorites.

CCB-168 - ASCII NonNegative Integer may or may not include + sign?
See <https://pds-jira.jpl.nasa.gov/browse/CCB-168>

Left this in a wait state until CCB-131 is resolved.

~ No. Wait. This is a new one.

Basically, the story is that as currently validated you are allowed to have a plus sign, but the standards don't allow that.

Question: Does anyone want to allow a plus sign instead of just a number?

Answer: (Silence)

~ Okay, will write it again.

We could report this as a bug because the validation is not performing correctly.

~ Someone wasn't clear that this is a bug.

Question: Does the SR say that a plus sign is not allowed?

Answer: SR says non-negative can have digits 0-9. Doesn't mention a sign. Validation is allowing a sign.

~ Sounds like a bug fix.

Some one is not convinced it needs to be prohibited, but if no one cares- fine. We'll just fix the validation.

Question: So can the author update jira?

Answer: Yes - may add it to the bug fix SCR. Will handle it. ****(Action Item - Lev)****

~ Bug fixes are typically sent to Steve to fix. No SCR necessary for this.

Author will withdraw the SCR. It is subsumed by another SCR - CCB-171. Will send an email to Ron.

Question: This is turning into a big thing. Can't the author just send Steve an email or make a verbal request?

Answer: Ron will wait to hear what happens.

Question: Anything else on this?

Answer: (Silence)

CCB-170 - Deprecate IEEE 754 NaN and Inf in Favor of PDS4 Special Constants
See <https://pds-jira.jpl.nasa.gov/browse/CCB-170>

We are waiting for CCB-131 to be resolved.

We need to make up our minds on special constants and then we can look at this.

CCB-171 - Split hardware-compatible ASCII numeric types from the unbounded
See <https://pds-jira.jpl.nasa.gov/browse/CCB-171>

Kind of hit on this already.

Question: Does anyone have anything to say?

Answer: Yes. This is a major change. This is one where we can't read ASCII integers or reals that are beyond the hardware type. Last time we discussed this there was confusion. The SCR has been modified. The solution was to define ASCII unbound and bound. The question was when to use them. The SCR was modified to address that. The bottom line is that it's hard to use if it's not compatible with the hardware type. The story is that the SCR has been updated and people should read it.

Question: Anyone?

Answer: Um. Not sure I understand the issue yet, but don't think it's a big deal. It just doubles the number of data types.

~ Others had that same concern back in October.

Someone likes splitting this into a long term and short term solutions. We can fix the unbounded with min and max. Looking at modeling symmetry, but not entirely up to date on this issue.

Question: If you were reading the data and some of the numbers exceeded the floating point - how would you read it?

Answer: Would read what python allows for max precision and throw the rest out. Engineering versus science mind.

Another Question: What if you were interested in the last numbers?

Answer: Someone would look at the file and make decisions based on that. Would write a program to pick out the most significant figures.

There's a problem when you have a standard where you can't read the data.

Question: Shouldn't the data be readable?

Answer: The standards are written for providers so they don't provide exotic data.

~ Someone agrees.

~ Scientists would write a program.

~ We have a lot of users who are looking for tools. They don't have funding to write new programs. We need to consider user needs, not just provider needs.

Question: What percentage of data does this affect?

Answer: Common data types. Would need extra work to make sure people can read them.

~ That extra work takes a lot of time and memory.

Another Question: Data? Not in a label?

Answer: Correct, but sort of both because modifying types.

~ On the data side, when we started development we chose ISO1179 as the basis. We knew we needed certain precision in data types. It was considered. The question becomes do we do this in general or to specific data types or allow this in the metadata. Seems like we are overriding the data types.

~ Binary types inherently include the precision.

~ Thought this is regarding significant digits and values. Thought we were talking about what DPs thinks is significant.

~ For that would have to use field format. No limit on what they can put in that. Not exactly what we are trying to address. Trying to address precision of data types. The question is if they specify a number that exceeds the precision - how we would handle that.

~ So, in some cases format is constraining but not for ASCII, so we need to provide precision some way.

~ Binary provides it by default. Seems ASCII should too.

Question: It seems like we are focused on ASCII Real data in product - we're not saying what we allow. What's the proposal?

Answer: Essentially a double.

~ So overloading the definition with the machine.

~ But also want to add bound and unbound.

~ Someone was trying to understand the bounded problem. Need ASCII Real bounded. Can write that.

Someone would like to hear from IMG or GEO. Given that we have limited time and budget - want the big data people to weigh in.

~ SBN is impacted by Julian dates.

~ IMG understands the problem, could maybe add in a bounded version so we don't have to change anything from the past. Not sure how much this affects us.

~ GEO - not sure we've ever run into this as a problem.

Don't think we have a problem where people have already used this in PDS4 yet. Regarding GEO's comment, not sure, won't claim it's a common thing. Julian dates are not super common. But we shouldn't have a standard that isn't readable by computers.

Question: Someone is looking at a Navy PDF of Julian dates. They make claims about precision. Will send it to the group to look at. The question is if we go to all this work, do we care what we might lose? Are we solving a problem?

Answer and another question: In terms of Julian dates, Mike A'Hearn says this is a problem. Others too. Multiple scientists say this is important. The problem is that a chunk of the standard is unreadable. In my mind that makes the whole standard poor. Am I really the only one who thinks that matters?

Answer: Adding two datatypes doesn't seem like a huge effort. It's been well defined in the SCR.

~ Someone agrees. Not sure why this is such a big deal. Seems like a simple solution.

Question: Is the SCR ready for a TA?

Answer: Author thinks so. Minor changes were asked for last time. They were made.

~ Someone commented on it in November.

~ We would need three new data types.

Question: What's the third one?

Answer: ASCII unbounded non negative integer.

~ Should change all the places this could come up. No reason to be inconsistent.

- ~ It is consistent. Didn't want three types - seems excessive for a rare case that might come up.
- ~ Rare versus extremely rare. Update the SCR and people can comment.
- ~ It's already been updated.

We can vote next time - after it's TAed.

It will be TAed. Everyone look at it so we can vote. ****(Action Item- Everyone)**** ~ Put comments in before it's TAed.

Question: Anything else?

Answer: (Silence.)

PDS4 equivalent for the PDS3 keyword SOURCE PRODUCT ID

IMG is not dead set on how we define source product ID. IMG needed it for InSight - defined processing history for it. Just needed a way to define this. Someone said they could use references. That's fine, thought it made sense in processing history.

Rings uses them in separate file and has reference to point to them, but no appropriate type for reference type.

~ Someone thought we had one for PDS4.

~ That's the problem - if it's a PDS3 file there's no easy way to reference it. There was a way to point to a file in PDS3, but it's not using a LID.

People are confused.

There was a proposed solution, but not in an SCR. Will look for it, but it was basically saying to use reference list. Internal reference.

Question: How does that point without a LID?

Answer: Reference list allows both internal and external.

~ GEO was going to use external.

Another Question: How does that work?

Answer: Doesn't require a LID. Need a structure that allows a reference that isn't a LID. We might need to add something.

~ For PDS3 the requirement was that DSID and product ID be unique.

~ So we need to add a PDS3 type of reference in addition to the internal and external.

~ Someone agrees.

~ Seems ugly.

~ Yes - but that's cause of the names internal and external.

~ Someone would prefer modifying the references we have. Less ugly.

Concern about messing with internal reference. Open to discussion. Processing history in the IMG LDD seems like it is of interest to other nodes.

~ Maybe. Not entirely sure.

~ Not sure we should generalize this.

~ This is broader than just IMG.

~ Agreement. IMG uses it for raw versus derived too. Not sure what that would mean.

Question: If this is processing history for IMG they can do anything they want. Are we being asked for a new keyword?

Answer: Not just IMG wants it.

~ We need to look at the provenance models. Processing history is very common.

Question: Was this specified by history in PDS3?

Answer: It was done many ways in PDS3.

~ Was source product ID.

~ I mean the processing, there was a history thing.

(Anne has joined the call)

History was only used for cubes.

~ IMG has processing history which was supposed to be computer parseable. Trying to decide how to include.

Question: Isn't it in the IMG LDD?

Answer: Yes.

~ Thinking of moving it up.

~ May be a need to move this.

Don't get hung up on processing history to solve the request for a source product reference.

~ Okay. And a second issue is source product ID. Seemed like it could be used but we need something more than we have.

~ Yes, but PDS3 data sets. Not sure it won't work. They have unique IDs. Would use the DSID and citation information and/or DSID and product ID.

~ Someone likes that.

Question: Reference text?

Answer: The citation.

Another Question: Did I miss a keyword?

Answer: No. We decided how to cite PDS3 data sets. Do that. Bibliographic.

Another Question: Where do you cite the product?

Answer: Append it to the DSID.

~ Having trouble visualizing it.

~ SBN uses this.

~ Someone likes the direction this is moving.

The point of the bibliographic citation is that you assume users will be able to find the products. Need to make sure future users would be able to find PDS3 data.

~ We aren't throwing PDS3 away.

~ Still not sure how well it works.

~ Certainly something to consider. Maybe for the Roadmap.

Question: Can SBN provide an example?

Answer: Ask Carol at PSI.

~ That may cover source product ID.

Comment - the issue of unique IDs across PDS3 - making them references may not get us all the way there. Fuzzy, but we can make them more specific.

Question: Anyone else?

Answer: ****Action Item - Lev**** will send an example if he can find one.

Question: Do people want processing history added to the model? Willing to discuss that.

Answer: Yes. Was expecting that or provenance to be part of the model.

~ A few years ago work was done on provenance class - didn't get far.

~ Not surprising. Processing history will be very different from data type to data type. We have a modification history class already.

Question: Is that to track changes to product content or to the label?

Answer: It's for whatever you need it for.

~ Someone uses it for products.

~ Product is data plus the label.

~ Someone disagrees. They are separate. Data object, description object and identification object. But concedes that modification history is supposed to cover everything.

~ Unclear if restricted to history, but wouldn't invent a new class.

~ Mostly description. Not computer parseable.

~ IMG uses processing history text.

Question: Is there already an implementation for IMG? Could it be used for shape model?

Answer: Very generic.

Another Question: Can it be sent to the group?

Answer: Yes, but still need to decide if useful to the whole group. This was done for IMG. **** (Action Item - Jordan)****

Question: Are we done? Anyone else? We are past time. Will do product update next time we meet and Steve wants to discuss composite structure. Anyone else? Two weeks?

Answer: (Silence)