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

Sent: Monday, October 31, 2016 11:06 AM

To: pds4ddwg

Subject:Notes from PDS DDWG 2016 10 27

---

title: DDWG Notes 2016-10-27

layout: default

date: 2016-10-27

October 28, 2016

Notes by Debra Kazden

Known Attendees:

E. Guinness, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, S. McLaughlin, 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 - October 26, 2016

CCB/SCR Statuses:

-- None

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

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

-- CCB-166: Deprecate bit mask from IM for new sample bits attribute (J.Padams) - please review SCR as has been updated

    -- 20160901: Open & Under DDWG review

        -- See comments from R.Simpson & Lev

    -- 20160829: email from R.Simpson to Jordan on suggested changes

    -- 20160908: many emails between Jordan, R.Simpson, and Lev

    -- 20160929: Jordan to re-write the SCR to only deprecate <bit mask>

        -- all other issues will be handled in IMG LDD

-- CCB-162: Move <md5 checksum> from Object Statistics to Tagged Digital Object (R.Simpson) -- please review SCR for possible vote

    -- 20160622: Open & Under DDWG review

        -- emailed Steve to TA

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

    -- 20160630: C.Isbell & Jordan & Dick to resolve issue(s); then back to DDWG

    -- 20160706: email to C.Isbell & Jordan & Dick asking if issues within WG are settled

    -- 20160804: email to Jordan asking about status / progress

    -- 20160829: email from R.Simpson with suggested implementation

        -- back to DDWG for discussion

    -- 20160901: R.Simpson retitled & rewrote SCR

    -- 20160915: Steve TA'd; but TA has lien

        -- Steve & Simpson to resolve

        -- Steve to generate two implementations & present to DDWG for discussion

    -- 20160929: Steve provided the two implementations to DDWG (for review / comment)

        -- discussion postponed because provided too late for review by the team

    -- 20161017: Simpson revised SCR; ready to vote

 -- 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

(2) Email exchanges and other stuff

-- ASCII_REAL (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

DDWG Telecon

People continue to have difficulty with WebEx. One meeting attendee reported having trouble using Firefox, but being able to use Chrome.
~ The problems with WebEx were reported after the last telecon.

This should be an interesting meeting.

Steve has a meeting at 10:00. Will leave this telecon early.

CCB/SCR Statuses

Nothing happened. Yay for our team.

CCB-166 - Deprecate bit mask from IM See https://pds-jira.jpl.nasa.gov/browse/CCB-166

Jordan is back from a pretty great trip now. It's been a long time. The SCR was rewritten. Basically wants to remove bit mask. The ticket was updated before Jordan left on vacation. Thinks it's ready to go.

It has not been TAed. Can't do much. Everyone should look at it. Maybe it can be TAed and we can do an email vote - seems pretty simple at this point.

Question: Is this reasonable?

Answer: Yes.

Action Item - Steve will TA the SCR.

Cool.

CCB-162 - Move <md5 checksum> from Object Statistics to Byte Stream See https://pds-jira.jpl.nasa.gov/browse/CCB-162
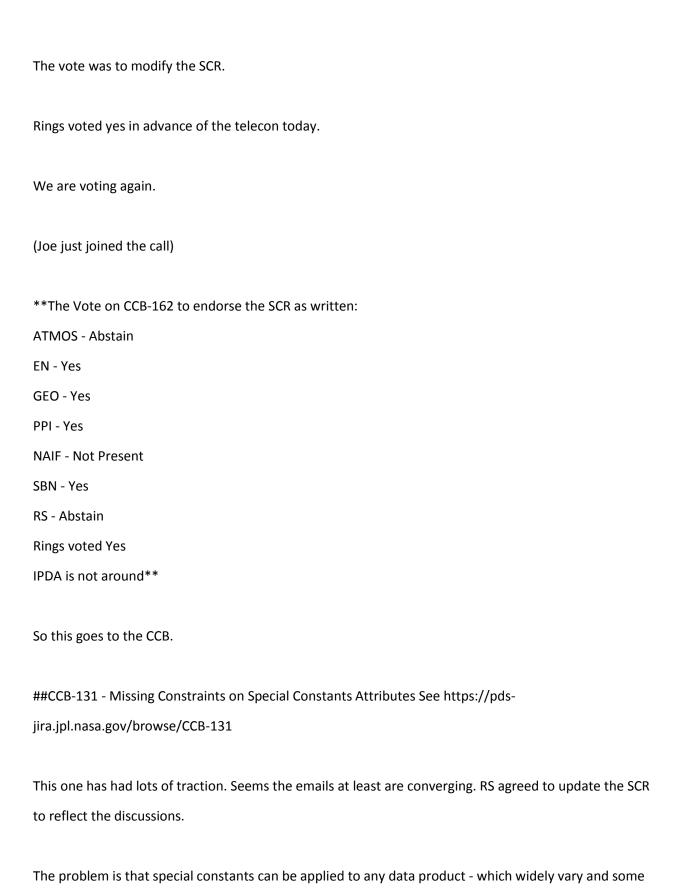
Hopefully, everyone has reviewed this.

Question: Is there any problem with the proposed implementation? We did vote on this and the SCR was updated.

Another Question: So the objections that NAIF had were settled since they voted yes on the thirteenth?

Answer: Yes.

~ Now it is backwards compatible.

~ It was changed so NAIF is okay with it, but it wasn't changed for that reason.

The vote was to modify the SCR.

Rings voted yes in advance of the telecon today.

We are voting again.

(Joe just joined the call)

**The Vote on CCB-162 to endorse the SCR as written:

ATMOS - Abstain

EN - Yes

GEO - Yes

PPI - Yes

NAIF - Not Present

SBN - Yes

RS - Abstain

Rings voted Yes

IPDA is not around**

So this goes to the CCB.

##CCB-131 - Missing Constraints on Special Constants Attributes See https://pds-jira.jpl.nasa.gov/browse/CCB-131

This one has had lots of traction. Seems the emails at least are converging. RS agreed to update the SCR to reflect the discussions.

The problem is that special constants can be applied to any data product - which widely vary and some

people find confusing. The proposal was to have special constant defined the way the parent is. If the data object is ASCII real, than the special constant would be ASCII real. There's an issue for binary. There was a proposal that all special constants have to be ASCII. Binary integers would be ASCII integers - there would be a potential problem for floating point numbers - can use ASCII real, but has to be in binary format. Whole numbers can be represented in binary - unless there are too many digits. Think that's it. Still some issues with if applied to stored values or if after scaling factors or offsets - probably our choice. Speaker would prefer stored values. Emails suggested we not rely on NANs or INFs - would prefer everything is explicit and defined in the label. We should define some values to tell users something is not a number. We need some agreement today and then we can update the SCR.

Someone wants to discuss NAN separately, but someone else is concerned because there are binary data with NANs and INFs in PDS4, from MAVEN. Guessing that was left alone in the label, not sure. NAN is not one value - data can have several NANs with several meanings. It's hard to tell them apart. It's not uncommon. Lots of programmers will tell you it's possible - if you allow NANs you need to use HEX byte patterns - hard with MSB and LSB - confusing. No suggestion, but wants to caution that it might be that DPs want it - having only strings might cause special constants to not be given in the label.

~ MAVEN is using 1.4 - they can keep doing that. This is for 1.8 and beyond.

~ The MAVEN data were corrected.

~ There were a lot of NANs in MAVEN.

~ Someone believes that was corrected.

~ There are some in the MAVEN PPI data. Some instruments have quite a few.

Someone is concerned over the documentation aspects of this. More than NAN as a convention.

~ Programs that produce NANs can put more information in to give information - the question is what is actually in the information load area of the NANs. Need to know if it's important and if it's documented.

~ That's one reason to put special constants.

~ The documentation is available through the IEEE standard where the binary standards come from.

~ That says the rest of the bits can hold information. May not all be zero - so, if you process using them and it's not documented - then we won't know if the archive is missing something.

We don't know. If someone is giving NANs we have to ask what's in the information load.

~ Good point, but this SCR is how you represent them. ASCII is a problem.

~ Yes, absolutely right.

Question: So, if things continue being allowed - if you have a missing constant in the data, but it's not in the label, does that violate the standard?

Another Question: How would you know that situation exists?

Answer: Reviewers could find it.

~ That would be a lien.

Another Question: Would DPs be forced to change it or could they leave it?

Answer: Believes it's not legal to include NANs in PDS4 binary data, so it would be invalid. The data would need to be corrected. If special constants isn't documented it will be hard to detect that.

~ It's not invalid now. Seems legal.

~ Scientifically it's not a good practice. Shouldn't use NANs and INFs.

~ Someone is surprised it's not specifically prohibited.

~ Someone else thought it was common sense.

~ No, it's not common sense. The standard is unclear - and we are only using part of it.

Question: Aside from NAN and INF, is the rest of this palatable for now? We can separate those out. We were supposed to define special constants for ASCII. That's the real issue.

Another Question:: So for binary we would use the closest ASCII equivalent?

Answer: Yes. The bit pattern should be an exact representation in IEEE754.

Another Question: The closest ASCII equivalent, not bit pattern - no bit patterns in the label, right?

Answer: Right, but in the data.

~ So it's in the documentation.

~ The value in the data file has to be an exact representation.

~ Documentation issue for PDS for standards/practices. We have to be able to tell what the numbers are - we can't validate that. It becomes a tricky issue.

~Probably can't validate in schema, not sure about schematron. A data validator could do it, depending

on who writes it.

~ Will need to assemble some requirements for this.

Question: Any knowledge of how scaling and offsets were used in the past?

Answer: **Action Item - Jordan** - will send email to Bob Deen to ask.

We can write this up and add that part later.

There are software concerns here too.

~ Let's see what Bob has to say. We can write it like the numbers come straight out of memory and can change it later. Will split off the NAN and INF stuff.

Email exchanges and other stuff - ASCII REAL

Question: From all of the email - not sure if we're converging. What do people think?

Answer: We have an active discussion. If the consensus is to limit ASCII real to 16 digits, fine.

~ It's magnitude and digits.

~ We don't run into problems until exponents.

~ There have been discussions about how to define this. Would like to postpone this now and continue in email.

~ Someone is fine postponing.

Sean (who is not participating in the telecon today) was asked by someone if the resolution of this issue would have any effect for him. He said no in terms of EN. No dog in this fight.

One person sent email asking if this is even worth discussing.

~ The answer depends on your discipline. It matters for radio science. They are getting 128 bit numbers.

Someone asked Sean about quads. No real answer. Can check with him again.

There should be a big impact on EN. Any time we have to handle beyond double precision it needs a big programming effort.

Sean was asked about transformations - no real answer.

It would be bad if we started ingesting really long numbers and we couldn't do anything with them.

~ No one wants to speak for Sean, but someone thinks there will be impacts, doesn't really know.

The decision is to bring this back in a few weeks.

~ Yes.

Next topic - It's on the MC agenda to get a resolution on packed data. There are three SCRs on hold pending that decision.

There's also the issue of if people can use an older IM. Seems packed data should be decided first, cause if no packed data than can use older IM and still have packed data.

~ That will be a problem for software.

~ Users will have to deal with the software.

In PDS3 people used TBtool. It was semi-decent. Better than how PDS4 does it now.

~ At the tech session Sean listed TBtool functionality - not sure what he will do to improve it.

_____

Question: Anything else?

Another Question: Cubs versus Indians?

Answer: (Silence)

Two weeks