Minutes v1.0
LSI-VC-9 Teleconference #1: CEOS ARD & Standardisation

Tuesday 14 April 2020

Participants

ESA:	Ferran Gascon, Ivan Petiteville
EC/JRC:	Zoltan Szantoi (Chair)
GA:	Adam Lewis, Andreia Siqueira, Medhavy Thankappan, Fuqin Li
IEEE:	Siri Jodha Khalsa, Chris Durell, Brandon Russell
JAXA:	Takeo Tadono, Ake Rosenqvist
KARI:	Chiho Kang, Daehoon Yoo
LSI-VC Sec:	Matt Steventon
RADI:	Jiao Weili
SIT Chair Team:	Stephen Ward
SEO:	Brian Killough
UK Catapult for UKSA:	Electra Panagoulia
USGS:	Steve Labahn, Jenn Lacey, Tim Stryker
WGCV:	Cindy Ong, Nigel Fox, Emma Woolliams
WGISS:	Rob Woodcock

The presentation slides compiled for this meeting are attached in Appendix A.

Introduction

Zoltan Szantoi (EC/JRC, LSI-VC Co-Lead) welcomed everyone to the first of four teleconferences that make up the virtual LSI-VC-9 meeting. This call is focused on CEOS ARD and formal standardisation processes. We will review the background of the CEOS ARD standardisation discussion and then hear a perspective on standardisation (pros and cons, options) from representatives from IEEE. This will include discussions on whether we proceed in this direction and the best approach for doing so.

Background

Matt Steventon (LSI-VC Sec) presented some background and the reasons behind LSI-VC's desire to discuss standardisation:

- CEOS ARD Product Family Specifications (PFS) are a standardisation of products but are not formal Standards.
- The 'if' and 'how' of whether the Specifications should be used as a basis for formal Standards is a topic of consideration under CEOS ARD Strategy Item 4.4.
- Standards may emerge as a key factor in CEOS engagement with the private sector.
- "Analysis Ready Data" has a different meaning to each user, and various groups are producing/working towards some form of "ARD" (e.g., Open Geospatial Consortium (OGC) Testbed-16). We have addressed this internally by focusing specifically on "CEOS ARD", but the issue remains that it is hard for users to understand exactly what "ARD" means.

- There is also the possibility of divergence with multiple definitions, which could undermine the goal of CARD4L – making data easier to use and increasing impact.
- Standardisation also has the potential to greatly increase the number of stakeholder groups engaged in CARD4L. CEOS Analysis Ready Data strategies are unlikely to succeed in the long term unless all key stakeholders are engaged.
- Some industry contacts have also flagged to us that community agreement on the definition of ARD is a necessity – going beyond just a written definition and including sample products and processes.

Discussion

- Ake Rosenqvist (JAXA) questioned whether anyone has the mandate to define ARD. In CEOS we can only define CEOS ARD. He noted that discussions with industry in WGCV revealed that companies mostly have their own definitions of ARD, and they tend to be agnostic on the exact form – as long as the data is ready for their particular use case. Ake suggested sticking to CEOS ARD, which in his eyes is the only thing we have authority to define.
- Ivan agreed with Ake and added that there has been limited interaction with end users to date in the CARD4L effort. Before considering any expansion of the definition effort we should broadly consult with users and continue to iterate based on their feedback and requirements. Any move towards standardisation should be driven by user need.
- Rob Woodcock (WGISS) suggested that whilst both of these points are valid, we could spend a lot of time trying to define ARD for everyone and discussing mandates and external user requirements. He encouraged LSI-VC to remember that the VC is composed of experts that are invested stakeholders in the use of EO data to solve issues related to the SDGs, climate change, etc. He suggests LSI-VC skip a few steps and instead consider whether CARD4L could be used as a useful springboard for gathering like-minded users together to make progress on these key goals. While there may not be a global mandate for CEOS to define a standard, there is a need for one, and if it were to emerge from CEOS in response to CEOS needs then that would be fine.
- Steve Labahn (USGS, LSI-VC Co-Lead) responded that our work has gone beyond the LSI-VC – we have engaged with the broader user community through the science teams, etc., that have been consulted in the development of the PFS.
- Adam Lewis (GA, LSI-VC Co-Lead) agreed that the process to date has been broadly consultative. He supported the comments of Rob Woodcock regarding the expertise and investment of those who have been engaged so far and noted that the work had proceeded over a period of 5 years, driven by individuals volunteering their time.
- Siri Jodha Khalsa (IEEE) suggested that the purpose of today's conversation is whether to use CARD4L as a springboard for a broader community consensus process for arriving at a more widely agreed definition of ARD and associated specifications.

Open Geospatial Consortium (OGC) Testbed-16

Zoltan Szantoi (EC/JRC, LSI-VC Co-Lead) reported his findings from a recent teleconference he joined on the OGC Testbed-16 initiative:

- Testbed-16 will undertake research and development in real-world environments to further advance technologies falling under the following threads: Earth Observation Clouds, Data Integration & Analytics (including ARD), and Modeling and Packaging.

- The ARD Testbed is managed by the University of Barcelona's Centre de Recerca Ecològica i Aplicacions Forestals (CREAF). CREAF also runs the Catalan Data Cube, which is built on Open Data Cube foundations.
- The lead likes LSI-VC's approach (general description and PFSs) but they are also interested beyond our domain (i.e., *in situ*) as well as descriptions/PFS for Lidar, and Oceanic and Atmospheric descriptors/PFS. OGC were keen on the idea of using CARD4L documentation (PFS, etc.) as a basis.
- OGC has many private companies as members, which could be useful for LSI-VC given our ambition to increase engagement with industry.
- There is the opportunity for others in LSI-VC/CEOS to become involved in the OGC Testbed-16 activity. Anyone interested can contact Zoltan for details.

Discussion

- Siri Jodha Khalsa (IEEE) suggested that it would be helpful to distinguish ARD (data products) from Analysis Optimized Data Stores (AODS). The PFS don't discuss how ARD is served/accessed.
- Rob Woodcock (WGISS) agreed this separation of concerns is important. CARD4L is a content standard – referring to interoperability between measurements (or corrected measures). Discovery, access and analytics are other forms of interoperability. The draft CEOS Interoperability Terminology Report attempts to describe these separations and their relationships. This is important to WGISS as the discovery and access will be impacted both by ARD and Cloud access and significant duplication of CEOS sourced data by third parties on the Cloud. New discovery questions will need answers, e.g.: "Which Landsat archive is closest to me, Cloud A or Cloud B and are they the same? Can that archive be used for harmonisation with Sentinel archive on Cloud C?".
- Siri Jodha noted that the focus of Testbed-16 is around standardising how data is served, rather than prepared.
- Chris Durell (IEEE) commented that CARD4L is a good standard to build upon. Industry's complaint is that it cannot meet these stringent requirements and is asking for another level, or levels, that are derived from CARD4L (or its like) but represent a lower but progressive path that industry can aspire to meet. Rob Woodcock (WGISS) agreed.
- Cindy Ong (WGCV) recalled from the ESA VH-RODA workshop last year that some of the comments were also that there are areas of CARD4L that are not quantitative enough.
- Adam Lewis (GA, LSI-VC Co-Lead) noted that Chris' point gives some context also for our broader engagement with industry. The progression of standards is only one of the things we are doing. We are also planning a workshop with industry in Europe (previously planned for September 2020, new date TBD).
- Ferran Gascon (ESA) supported focusing on our work on CARD4L and adapting it to address this concrete requirement from industry (different levels of compliance). He sees good prospects for progress on industry engagement along these lines, rather than the standardisation approach.
- Nigel Fox (WGCV IVOS) noted that IVOS instigated a number of teleconferences following on from the ESA VH-RODA workshop with industry, USGS, other CEOS Members. The team came to the conclusion that it will look to start building example addons to evolve CARD4L to cover different sub-classifications of ARD between threshold and target. These classifications would also be grouped by application type.

- Rob Woodcock (WGISS) agreed with Nigel on the need for gradation. Experience supports this as some resistance will be quite vocal, preventing broad adoption, if for example publishing a quality standard is embarrassing for the provider of the data. They may well be producing quality data, but not in a form that can be readily published and accessible. Gradations provide opportunities for evolutions of systems and processes.
- Ake Rosenqvist (JAXA) noted that this is similar to the 'partial compliance' discussion from earlier meetings. He is supportive of having some means for partial compliance, as long as it is supported by the appropriate documentation.
- Adam Lewis (GA, LSI-VC Co-Lead) replied that Nigel's intervention is interesting but challenging. The foundation for the CEOS ARD concept is estimating a property of Earth's surface, so to go below that is to be doing something different. We also need to be thoughtful about whether producers can afford, or users really want, a spectrum of CEOS ARD products. Adam and Steve supported the need for accuracy measures – this is necessary and will be transformational.
- Adam agrees with Ake that we should allow providers to be transparent about which parts of the specifications they meet, and which not. However, in his experience only a specific user can decide if the data are fit for purpose. Nigel agreed that only the user can assess suitability, but it can be made easy for them by providing granularity and ease of selection, and in some cases guidance on what others find useful.
- Rob Woodcock (WGISS) suggested that the proposed granularity might need to be part of the query used in a discovery process, and this is something for WGISS to track when updating CEOS discovery systems to accommodate CEOS ARD.

IEEE Perspective on Standardisation

- Siri Jodha Khalsa (IEEE) presented slides 8 to 15 (see Appendix A).
- Adam Lewis (GA, LSI-VC Co-Lead) noted that we need some mechanisms to engage with the broader community. We now have approval to engage with industry following endorsement of the CEOS ARD Strategy. The standardisation approach offers another opportunity to discuss the CARD4L specifications with the wider community.
- Zoltan Szantoi (EC/JRC, LSI-VC Co-Lead) noted that there are three options: OGC, ISO, and IEEE. He asked Siri which he thinks would be the best way forward. Siri suggested that OGC's focus in general is not so much about the form of data, rather how it is accessed, used, served, etc.

Discussion

- Ake Rosenqvist (JAXA) thinks that engaging with standards organisations could increase feedback on the PFS. This could also be a good opportunity to reach a consensus on related terminology.
- Siri Jodha Khalsa (IEEE) suggested that working with OGC as a start will provide lessons, feedback, etc., however, OGC is not going to produce an ARD 'Standard'. OGC could be a good discussion forum.
- Siri noted that standards bodies will only meet a certain number of times per year to consider new submissions, so getting started early is important.
- Zoltan Szantoi (EC/JRC, LSI-VC Co-Lead) asked whether we need a further mandate from CEOS to progress down these routes. Adam Lewis (GA, LSI-VC Co-Lead) suggested this would likely be

necessary (at a SIT Technical Workshop, for example) before making a commitment, however there's nothing stopping us from exploring the idea, given the approval of the CEOS ARD Strategy.

- Chris Durell (IEEE) noted that the ISO and IEEE processes uniquely bring a level of industry recognition and invitation to participate. He expects that there would be a good response from industry should we go down this route. These options do not preclude discussions in OGC, etc.
- Adam noted that ISO's requirement for a national representative to support the application complicates that option.
- Siri reported that the IEEE process would take place through the IEEE GRSS Standards for EO Technical Committee, which he leads. IGARSS meetings are used as opportunities to convene the committee, but the process can be initiated at any time through the submission of a Project Authorization Request (PAR). CARD4L is developed enough to begin the PAR, but LSI-VC would need to decide the scope and form (e.g., there could be a PAR for each PFS, or they could be subsections of an encompassing PAR which can be added as needed). Chris shared the PAR for P4001 Hyperspectral as an example.
- It was noted that the resulting IEEE standard would need to be purchased, and CEOS would lose 'ownership' of the specifications. The cost could be a deterrent to the groups of users we are trying to serve.
- Ivan Petiteville (ESA) urged that the process must be driven by end and intermediate users' needs, which data producers will have to assess and consider the feasibility of implementing.
- Electra Panagoulia (UK Catapult) asked whether there might be an impact on compliance due to the use of proprietary software to generate ARD. Zolti noted that the CARD4L PFS are not prescriptive – as long as the outputs are within the specifications the data can be CARD4L compliant. Emma Woolliams (WGCV) discourages writing specific software into any standardisation – it is much better to require openness about what is used, with good metadata and regular comparisons carried out.
- Steve Labahn (USGS, LSI-VC Co-Lead) asked whether we (the land part of CEOS ARD) could move along in this direction independently of the CEOS ARD for oceans and atmosphere discussions taking place in CEOS. Adam suggested we can work to understand the process and take it back to CEOS for discussion before committing to anything. Adam doesn't think land should proceed independent of the other domains.
- Steve asked whether others in CEOS, maybe SST-VC, have already gone down the standardisation path. Siri Jodha has communicated with Ed Armstrong (NASA/JPL, SST-VC Co-Lead) and can confirm that SST-VC has not been through the standards process. Nigel Fox (WGCV) also confirmed they have specifications but have not entered any formal standardisation process.
- Adam noted that the problem with the OGC Testbed-16 approach is that the ARD discussion becomes tangled with the Data Cube discussion. Out of the ISO and IEEE options, IEEE seems the path of least resistance (no need for a national sponsor). We also need to understand the resources available to undertake such an effort.
- Siri communicated the following from George Percivall (OGC CTO and Chief Engineer): *"As the CEOS CARD4L specs describe the calibration and validation for remote sensing, e.g., radiometric calibration, it makes sense that the CARD4L specs are more relevant to IEEE GRSS than OGC standardization. It would be good if the CARD4L specs would reference OGC specs when relevant."*

- Cindy Ong (WGCV) noted that ISO already has a series of standards on remote sensing cal/val (19159 series). Siri Jodha said CARD4L could define a profile of 19159 standards. The optical part of 19159 may be due for review soon, which could be timely. Cindy noted these processes are very long and time consuming.
- Nigel Fox (WGCV IVOS) questioned why there is the objective to venture into the formal standards process. He noted that no datasets have yet been classified as CARD4L, but we are effectively already discussing reopening the definition of specifications. Adam suggested that we are exploring this as a possible way to enhance the reach of CARD4L in support of our initial goals (increased uptake and impact of EO data). Noting the loss of ownership/control associated with the standards process, Nigel advocated for continuing to develop the specifications under CEOS and making more effort to bring external users into the discussion (including from industry), as this seems more manageable.
- Siri noted that IEEE working groups are either entity-based or individual-based. If CEOS were to lead this as an entity, it would need to become a corporate member of IEEE. If individual-based CEOS couldn't contribute as a collective group. Either way, CEOS would also lose its control/branding of CARD4L.
- Rob Woodcock (WGISS) suggested that one of the advantages of achieving an official standard is that groups, especially government groups, that are hesitant on adoption because of the costs involved, tend to have an extra reason to adopt something because it is a standard. The major disadvantage is loss of control and the effort required. OGC was formed because of the latter issue – ISO standards don't keep pace with current innovation rates. For CEOS, he suspects timeliness and relevance to CEOS purposes is paramount, probably more so than the international standards.
- Ferran Gascon (ESA) suggested that connecting with downstream users to get their feedback is a necessary step before committing to any position regarding standards (industry in particular).
- Cindy Ong (WGCV) noted that sometimes “pseudo-standards” which fill a gap are just as valued and used (e.g., the WGCV Protocols).

Closing

- Steve Labahn (USGS, LSI-VC Co-Lead) noted that the next LSI-VC-9 call on the 28th of April is about industry engagement, and this may provide the missing pieces that we need to make an educated decision (during the fourth and final LSI-VC-9 teleconference) on where we want to go regarding standardisation.
- Zoltan Szantoi (EC/JRC, LSI-VC Co-Lead) thanked everyone for their attendance and very valuable contributions to the discussion. He thanked Siri Jodha Khalsa (IEEE) in particular for his guest presentation.
- LSI-VC-9 Teleconference #2: Industry & CEOS ARD will be held on April 28, 07:00 – 10:00 US East (other local times) [Presentation].

Appendix A: Meeting Presentation Slides

CEOS ARD & Standardisation

LSI-VC-9 Teleconference #1

1

Overview

- Background of the CEOS ARD standardisation topic, including industry feedback.
- IEEE perspective on standardisation (pros and cons, options).
- Standardisation discussion

2

Background

LSI-VC Leads

3

CEOS ARD Standardisation Background

- CEOS ARD Specifications are a standardisation of products, but are not formal Standards.
- The 'if' and 'how' of whether the Specifications should be more formalised as Standards is a topic of consideration under CEOS ARD Strategy item 4.4:

4.4: Engagement with standards organisations	CEOS Agencies, SIT Chair	Q1 2021 Onwards
CEOS to progress CARD4L discussions with standards organisations (e.g., OGC) to explore whether CEOS ARD specifications might be used as the basis for broader, official community standards, and, to ensure that CEOS work is recognised by others including the data research community.		

- Standards may emerge as a key factor in CEOS engagement with the private sector.

4

Why CEOS ARD & Standardisation?

- "Analysis Ready Data" has a different meaning to everyone
- Various groups producing/working toward some form of "ARD"
  - The Open Geospatial Consortium (OGC) is investigating Analysis Ready Data, at the request of an undisclosed patron.
- We have addressed this internally by focusing specifically on "CEOS ARD"
- But the issue remains: hard for users to understand exactly what "ARD" means
- Possibility for divergence which might undermine efforts?
- Engage with standards processes to support independent formalisation of "ARD" ?
- Include all stakeholder groups. CEOS Analysis Ready Data strategies will not succeed in the long term unless all key stakeholders are engaged.

5

Industry Feedback

- Need a joint acceptance of what ARD is – going beyond just a written definition and including sample products and processes.
  - Sample products and documented processes that could be used to confirm products are accurate and CARD4L consistent would be helpful.
  - Support from CEOS with cal/val, e.g., access to site data

6

OGC Testbed-16

Testbed-16 will undertake research & development in real-world environments to further advance technologies falling under the following 3 thread

- Earth Observation Clouds
- Data Integration & Analytics ->ARD
- Modeling And Packaging

The ARD TESTBED is managed by University of Barcelon's CREAF. The lead likes LSI-VC's line (general description, and PFSs), but they are also interested beyond our domain (i.e. in-situ), as well as description/PFS for Lidar, and Ocean and Atmospheric descriptors/PFS

- CREAF runs the Catalan Data Cube built on the AUS DC

OGC has many private companies as members, the involvement of them is important - or the outreach to them through OGC

7

IEEE Perspective on Standardisation

Siri Jodha Khalsa (IEEE GRSS Standards for EO Technical Committee)

8

IEEE-GRSS Perspective on Standardisation

- Pros
  - Increases buy-in from the broader stakeholder community, reduces confusion
  - Establishes consistency in terminology, concepts and procedures
  - Formalizes compliance criteria
- Cons
  - Requires volunteer effort
  - Must follow procedures to ensure openness and fairness
- Options
  - IEEE Standards Association
  - ISO/TC211
  - Open Geospatial Organization

9

Principles of Open Standards Development

- Cannot be controlled by any single person or entity with any vested interests;
- Evolution and management carried out in a transparent process open to all interested parties;
- Standard is platform independent, vendor neutral and usable for multiple implementations;
- Openly published (including availability of specifications and supporting material); and
- Approved through due process by rough consensus among participants.

10

Paths to Standardization - IEEE and ISO

- Develop proposal
  - IEEE: Project Authorization Request (PAR)
  - ISO: New Work Item Proposal (NWIP)
- Marshall support
  - IEEE: PAR reviewed by AudCom, Voted on by NesCom
  - ISO: participation by 5 member countries (National Standards Institutes) required
- Form working group
  - IEEE: Broadcasts Call for Participation
- Draft standard
- Ballot
  - IEEE: draft voted on by WG, then submitted to RevCom, final approval by SASB
  - ISO: Approval by two-thirds of member countries required
- Five-year review cycle

11

IEEE-SA Options

- Standard
  - A document with mandatory requirements
  - Characterized by the use of the verb "shall"
- Recommended Practice
  - A document in which preferred procedures and positions are presented
  - Characterized by the use of the verb "should"
- Guide
  - A document in which alternative approaches to good practice are suggested, but no clear-cut recommendations are made
  - Characterized by the use of the verb "may"

12

ISO/TC211 Options

- International Standard
- Technical Specification
  - For when the subject in question is still under development or where required support for an International Standard is lacking
- Technical Report
  - For material related to development of International Standards
  - Entirely informative, i.e. non-normative
- Publicly Available Specification
  - An intermediate specification, published prior to the development of an International Standard
  - Is normative but does not fulfill the requirements for a standard

13

OGC Options

- Two standards development tracks: Full standard or Community standard
- Three different Member organizations endorse submission of a proposal
- Lead for the submission team must be a Voting Member in OGC
- For a candidate Community standard, there must be evidence of implementation and commitment to support the implementation
- Other types of outputs
  - Discussion Papers
  - Engineering Reports
  - Best Practice Documents
  - Community Practice Documents
- BPs or CPs may transition to full standards through the OGC process

14

Resources

IEEE-SA

https://standards.ieee.org/develop/index.html

ISO

https://www.iso.org/stages-and-resources-for-standards-development.html

OGC

https://www.ogc.org/ogc/policies

15

Discussion: CEOS ARD Standardisation

16

Seed Questions / Notes

Should we seek to formally standardise?

- Would this path support CEOS objectives with ARD?
- Do we agree with the Principles of Open Standards?
- Are the CARD4L PFSs mature enough to enter a standards organization process? Or too mature?
- Are the other VCs on-board from a potential CARD4A and/or CARD4O perspective?
  - Would this help to build consensus or do we need to have full consensus in advance?
- Can we explore without commiting?

17

Seed Questions / Notes

If so, Which path? What are the challenges and opportunities?

- All offer “paths to maturity” from an idea through to a standard
- ISO requires national agency support which is likely to be difficult
- OGC can be initiated by a voting member (such as Geoscience Australia) ..
- Where would the resources come from?
- Which path would best recognise the work already done?
- …. ?

18