



EarthServer-2

654367 EarthServer-2
Agile Analytics on Big Data Cubes
H2020 INFRA-1-2014 topic (7)

Deliverable: D 5.3

Standardization Report - Consolidated

Nature of deliverable: R
Dissemination level: PU
Scheduled Delivery Data: 30th April 2018
Status: FINAL

Prepared by: Peter Baumann
Verified By Angelo Pio Rossi
Approved by: Peter Baumann

Table of Contents

Table of Contents	2
Executive Summary	3
List of acronyms and abbreviations	4
1 Introduction.....	5
2 Achievements per WP.....	5
2.1 OGC (T5.1).....	5
2.2 ISO (T5.2)	8
2.2.1 ISO Coverage Standards.....	8
2.2.2 ISO SQL/MDA.....	8
2.2.3 Work Beyond EarthServer-2	9
2.3 INSPIRE (T5.3).....	9
2.4 RDA (T5.4)	9
3 KPIs.....	10
4 Conclusion	12

Executive Summary

EarthServer-2 WP5 addresses “datacube” standardization in relevant bodies and, additionally, outreach work in the Research Data Alliance (RDA).

In this report, we summarize progress made in the reporting period, month 19 (following mid-term review) through 36 (end of the project). During this one and a half year period EarthServer-2 standardization work has led continued producing significant and internationally visible results in OGC, ISO, and INSPIRE. Also in RDA substantial and visible results have been accomplished.

Where standardization work has not yet led to adoption within the lifetime of EarthServer-2, work will continue after the project to ensure results will ultimately be available.

List of acronyms and abbreviations

CIS	Coverage Information Schema
ESA	European Space Agency
GML	Geography Markup Language
INSPIRE	European legal framework for a common spatial data infrastructure
ISO	Standardization Organization
MDA	Multi-Dimensional Arrays
OGC	Open Geospatial Consortium
RDA	Research Data Alliance
SQL	Structured Query Language
URL	Uniform Resource Locator
WCS	Web Coverage Service
WCPS	Web Coverage Processing Service

1 Introduction

EarthServer-2 WP5 addresses “datacube” standardization in relevant bodies and, additionally, outreach work in the Research Data Alliance (RDA).

In this report, we summarize progress made in the reporting period from M3 (June 2017) through M4 (May 2018). In this reporting period, the PI and his team participated to all relevant meetings either physically or via telecons, leading to distinct advances in standardization, and also led to a significant datacube technology report published in the reporting period.

2 Achievements per WP

2.1 OGC (T5.1)

Standards work. In the course of the project, the following standards have been suggested:

- Coverage Implementation Schema (CIS) 1.1 – adopted
- Web Coverage Service (WCS) 2.1 – adoption vote ending 2018-may-20

OGC Architectural Board (OAB). The OAB conducts fortnightly telecons (in places even weekly), plus both an open and a closed session at every OGC Technical Committee meeting. Peter Baumann being OAB member has participated regularly. Topics discussed include sanity checks for specifications being in the standards adoption pipeline; technology trends; liaisons; and more.

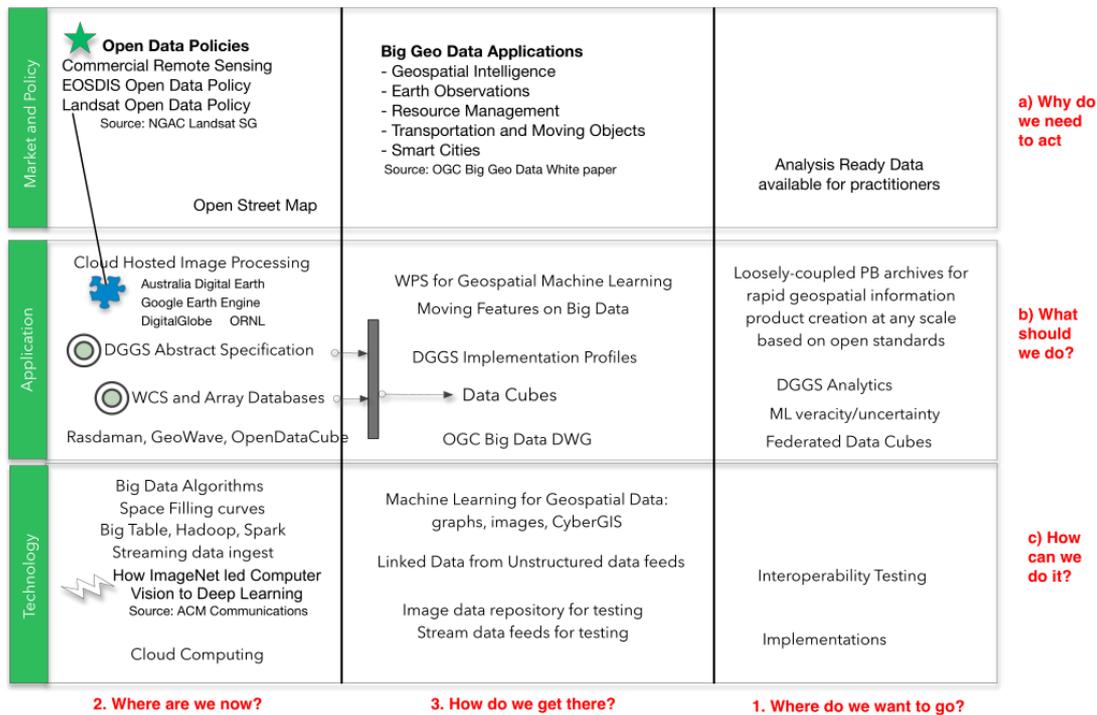
While coverages are well on the radar of OGC they have gained particular thrust when it turned out that they offer a convenient data and service model for datacubes. Since then, the datacube topic also has found its way into OAB’s strategic discussion. Figure 1 shows the current (internal, unpublished) draft of a technology matrix. Note that Datacubes are at the center, WCS and Array Databases are listed as state of the art, expressly mentioning rasdaman.

TC meetings. Peter Baumann has attended all quarterly OGC Technical Committee (TC) meetings, either directly or through telecom. Presence at the TC meetings is important as opinion shaping and voting on candidate standards happens there. Being co-chair of the WCS.SWG, Coverages.DWG, and BigData.DWG, Peter Baumann also had to facilitate these meetings (his co-chair, Stephan Meissl from EOX, has rarely shown up at meetings over the last years). Additionally, liaison with relevant stakeholder WGs have to be maintained, such as MetOcean.DWG, ESS.DWG (Earth System Science), Temporal.DWG, CRS.DWG, and others.

Numerous presentations have been made at the TC meetings on topics of conceptual background of datacubes; specifications derived (such as CIS 1.1) and supported (such as the WCS MetOcean Application Profile), implementation aspects (i.e., new features of clients and servers developed in EarthServer-2), and services deployed (i.e.: the EarthServer-2 services). This has more and more brought datacubes to the attention of OC, and in its unofficial draft technology roadmap (Figure 1) OGC explicitly mentions WCS, Array Databases, and rasdaman in the context of datacubes. Also, the concept of federated datacubes appears here following presentation of EarthServer-2 ECMWF/NCI federation (including live demos).

Big Geo Data Analytics Technology Roadmap

Fri Nov 10 2017



LEGEND: Technology Enabler Market/Policy Stimulus Enabling Standard Application Achievement Dependency

Figure 1: OGC Draft Technology Roadmap

Datacube Working Group. During summer 2017 the charter for an OGC Datacube Domain Working Group has been prepared under the lead of rasdaman GmbH, supported by Jacobs University, NASA, European Union Satellite Center, Envitia/UK, and KEYW Corp./US. Goal of this envisaged working group should not be to establish standards – with the coverage standards suite OGC already has adequate datacube standards in place – but rather to establish a focal point of discussion and information on this cross-cutting theme. This is the summary of the charter:

The purpose of this Datacube Domain Working Group (Datacube.DWG) is to maintain an open forum for the discussion and presentation of datacubes in all their aspects – including, but not limited to, theory, practice, technology, “analysis ready” data organization, and standardization in the datacube context.

As a DWG, this WG will not establish RFC submissions, candidate standards, or revisions to existing OGC Standards. However, the DWG can develop change requests as document interoperability requirements for submission as work items to a SWG or to TC for consideration.

As datacubes are being addresses in a variety of bodies recently, Datacube.DWG will maintain close liaison with at least the following bodies: ISO; INSPIRE; W3C; RDA maintaining mutual information exchange. In particular, the group will contribute to the Research Data Alliance (RDA) Array Database Assessment Working Group (ADA:WG).

In September 2017, the OGC Coverages Domain Working Group unanimously moved to establish such a group. However, likely due to lobbying from competing groups in Australia and the US this motion has not been treated according to OGC's Rules & Procedures, but still is held back by OGC management at the time of this writing. Publication of the charter for RFC (Request for Comment), foreseen for 2017, has been envisaged meantime for some time "soon" in 2018.

Coordinate Reference System (CRS) resolver operation. A resolver is a Web service which, given a URL identifying a concept – in this case: CRSs – responds with the definition of the concept (in this case: a GML document describing the CRS). Technically, the CRS resolver implementation, named SECORE, is an open-source Java servlet based on an XML database system being part of (and maintained with) open-source *rasdaman community*. Since several years now JUB is operating OGC's resolver, and whenever OGC requests adding new CRSs JUB is doing the necessary updates.

Comparison to DoA. Relative to the DoA, achievement in this period is as follows:

- *co-chairing WCS.SWG, Coverages.DWG, BigData.DWG, Temporal. DWG:* WCS.SWG and Coerages.DWG continue to be led by JUB/RAS. Work in Temporal.DWG has been abandoned due to inefficiency of this group to produce concrete results. Datacube.DWG has been initiated, but is being blocked by Australian and US datacube stakeholders objecting to the European lead.
- *editor of new Big Data specifications and guiding them through the adoption process:* established CIS 1.1 (adopted), WCS 2.1 (under adoption), WCS-REST (failed).
- *guidance to other Big Data standards writers in OGC:*
 - GRIB2 Coverage Encoding 1.0 – adopted
 - MetOcean-WCS 1.0 – on hold since several years now, due to insufficient technical quality of the specification, as perceived by the MetOcean and WCS groups, and a lack of agreement on functionality among the stakeholders¹
 - EO-WCS 1.1 – adopted
- *extra work, not in DoA:*
 - OAB membership
 - Datacube.DWG proposed
 - CRS resolver operation

Summary. The overall goal of shaping technically sound standards that directly or indirectly relate to datacubes has been fully achieved – there is a remarkable footprint of impact through EarthServer-2 and Europe at large.

As this is a continuous tasked requiring permanent attendance, activities will continue beyond EarthServer-2: The OGC CRS resolver will be operated by Jacobs University;

¹ For this reason, MetOcean-WCS has not been implemented in WP3 (as originally foreseen in proposal). Rather, the core innovation of this specification, polygon/raster clipping, has been developed independently in a rigorous manner; this in turn will likely lead to a dedicated WCS Clipping Extension standard in the near future.

establishing and driving this Datacube.DWG will be pursued further, likewise leadership and active participation in the OGC standardization process.

2.2 ISO (T5.2)

EarthServer-2 is engaged in ISO in two distinct arenas. In TC211 WG6, rasdaman GmbH and Jacobs University co-shape the coverage standards, with high emphasis on aligning them with the corresponding OGC standards. In SC32 WG3, rasdaman GmbH has initiated and is driving an array extension to SQL. We inspect both below.

2.2.1 ISO Coverage Standards

TC 211 WG6 deals with geographic imagery, and their services. Already earlier it has been decided to revamp the outdated ISO coverage standard 19123 and replace it by

- ISO 19123-1 Abstract Coverage Model
- ISO 19123-2 Coverage Implementation Schema

Both standards will be aligned, with 19123-1 being the abstract conceptual model and 19123-2 being the “concrete”, interoperable, and implementable specification.

For 19123-2, a decision has been made to adopt OGC Coverage Implementation Schema (CIS), which is edited by Peter Baumann. Currently, OGC CIS 1.0 is under FDIS ballot as 19123-2; subsequently, adoption of OGC CIS 1.1 is planned. Peter Baumann is project leader.

For 19123-1, Peter Baumann has been addressed to write it (by 19123-1 project leader, Liping Di / George Mason University and Douglas O’Brien / convenor of WG6). There is a delay as Peter Baumann clarified that he cannot work on several specifications simultaneously, so 19123-2 must be finished before 19123-1 can start. Meantime, German DIN is sponsoring a project where RAS will write the specification and submit it via DIN to ISO, with OGC to be included for political reasons. OGC has announced that it plans to adopt 19123-2 as OGC Abstract Topic 6 update.

Comparison to DoA. Relative to the DoA, achievement in this period is as follows:

- *RAS will establish the OGC coverage standard as ISO standard:*
ISO 19123-1 is under FDIS ballot, 19123-1 writing has commenced.

2.2.2 ISO SQL/MDA

SC32 WG3 maintains SQL. Following an initiative by RAS, WG3 has accepted addition of n-D arrays to SQL and has tasked RAS – being member of German DIN after Jacobs decided to not become member due to the fees – to establish the specification. The result, about 10 pages of rigorous formal semantics specification, has been developed mainly by Dimitar Misev and is a core part of his PhD thesis at Jacobs University to be defended in May 2018. MDA integrates n-D array definition and manipulation seamlessly with relational tables.

Comparison to DoA. Relative to the DoA, achievement in this period is as follows:

- *RAS will continue work as editor of SQL/MDA:*
This has been driven to the point where ISO 9075 Part 15: MDA (Multi-Dimensional Arrays) is being released for FDIS vote.

2.2.3 Work Beyond EarthServer-2

Both RAS and JacobsU will continue standardization work in ISO until (at least) final adoption of both 19123-x and 9075-15.

2.3 INSPIRE (T5.3)

INSPIRE is a little involved because the status is different for coverage data and service specifications. As for services, INSPIRE has adopted OGC WCS 2.0 in December 2016 as a “Coverage Download Service”. This is precious because it aligns INSPIRE with OGC, hence allowing vanilla WCS implementations to be used in an INSPIRE context.

Situation is more involved with the coverage data model. Despite repeated warnings by Peter Baumann, INSPIRE in 2012 has adopted a coverage model which diverges from OGC / ISO, and which has some severe inconsistencies. Only recently, during INSPIRE implementation attempts, these have surfaced. Starting with joint seminars at the INSPIRE 2017 conference, Jordi Escriu, facilitator of the Thematic Cluster on Elevation an Orthoimagery, has teamed up with Peter Baumann to establish a consistent INSPIRE coverage definition, ideally based on OGC CIS 1.1.

During the INSPIRE 2017 conference, an unexpected liaison got initiated between the main (European) proponent of the Sensor Observation Service (SOS), Kathi Schleidt, and the EarthServer-2 WCS forces – after SOS and WCS have been competing for 5+ years in the INSPIRE arena. Reason for this new alliance is that SOS also attempts to deliver coverages as retrieval results, and confirmed inconsistencies in the current INSPIRE coverage definition which Peter Baumann, as OGC coverage expert and INSPIRE national delegate, had pointed out earlier.

Currently, the three proponents – Kathi Schleidt / AT, Jordi Escriu / ES, Peter Baumann / DE – are teaming up to convince a reluctant JRC about the value of CIS 1.1². A decision likely can be expected until summer 2018, before the annual INSPIRE conference in Fall where the three aforementioned experts will conduct an INSPIRE coverage training workshop.

Comparison to DoA. Relative to the DoA, achievement in this period is as follows:

- *JUB will participate in INSPIRE to establish OGC WCS as a Download Service for raster data:*
Accomplished earlier, currently work continues so that INSPIRE adopts the OC CIS 1.1 standard as underlying coverage model.

2.4 RDA (T5.4)

Research Data Alliance (RDA) is not doing standardization, strictly speaking. However, it is an important activity endorsed and supported by the European Commission, and a specific task in EarthServer-2.

Peter Baumann is co-chair of the Big Data Interest Group, the Geospatial Interest Group, and the Array Database Assessment Working Group (ADA:WG) – the latter chaired jointly by him and Kwo-Sen Kuo from Bayesics and NASA. This WG has

² See the statement by Kathi Schleidt on <http://myogc.org/go/coveragesDWG>

established a comprehensive overview on the state of the art on Array Databases and related array technology. This includes a landscape of relevant standards, tools, and operational services in the field, plus detailed functional comparison of 19 tools and comparative benchmarks of 4 tools (due to the massive heterogeneity of systems each extra tool imposes substantial extra adaptation effort in the benchmark suite, hence due to resource constraints not all 19 tools could be tested). The report has been presented at the 11th RDA Plenary in March, including a unanimous adoption vote by the ADA:WG. The document, constituting the most comprehensive survey available to date, is publicly available on the [ADA:WG wiki](#), with more than 1,100 accesses in the first month after its publication.

Comparison to DoA. Relative to the DoA, achievement in this period is as follows:

- *continue co-chairing the Big Data Analytics Interest Group and the Geospatial Interest Group:*
done continuously, and extended by establishing and leading the Array Database Assessment WG which has produced a significant outcome.
- *Peter Baumann has been suggested for the Technical Advisory Board (TAB):*
...but not been elected – hence, no activity could be accomplished.

3 KPIs

Table 1: KPI qualitative assessment for current reporting period.

KPI	Results	Achieved ?
The standards developed will increase prominence in standardization and advisory bodies.	JUB and RAS are firmly associated with datacube standardization in global perception, such as by ESA: “ <i>The RASDAMAN product is currently the world leading environment in this domain and the standard working horse for OGC standardisation on these innovative data access interfaces.</i> ” (Guenther Landgraf, Lead of Ground Infrastructure Engineering Department, ESA, 2017)	✓
Cross-fertilization between standardization activities will be exploited continuously.	JUB has developed the ISO 19123-2 spec from OGC CIS 1.0 written by JUB. RAS is developing ISO 19123-1, to be adopted as OGC Abstract Topic 6 subsequently. Both these work items are underway and being continued beyond EarthServer-2.	✓
A monograph will be written by the JUB and RAS teams explaining the in-practice use of the Big Geo Data standards OGC WCS and WCPS.	For SQL/MDA, a Technical Report has been written which is also in adoption by ISO. The books have not been written due to resource constraints, as standardization communication took	✗

Likely, the same will happen at some time on SQL/MDA.	substantial resources.	
Participate in standardization meetings and promote project findings	JUB has participated in every OGC TC meeting, RDA meeting, INSPIRE meeting, ISO TC211 meeting - usually in person, sometimes (travels too far or adverse schedule) by videoconferencing. Typically, JUB has facilitated more than one WG session at OGC meetings. EarthServer findings have been reported at each such meeting. RAS has participated physically in all ISO SQL working group meetings.	✓
Not in DoA, came up in the course of standardization work	<ul style="list-style-type: none"> OGC WCS REST interface (not adopted) Rewrite of ISO 19123 -> 19123-1 	

Table 2: KPI quantitative assessment for current reporting period.

KPI	Achievement
number of specifications, edited or contributed to by EarthServer members, submitted to a standardization body	<ul style="list-style-type: none"> OGC: 1 (CIS 1.1) ISO: 2 (SQL/MDA, after revisions still in adoption process; CIS 1.1 for 19123-2)
number of specifications, edited or contributed to by EarthServer members, adopted by a standardization body	<ul style="list-style-type: none"> OGC: 1 (CIS 1.1) ISO: 0 (CIS 1.1 in progress)
document level of specifications, edited or contributed to by EarthServer members (such as Discussion Paper vs Best Practice Paper vs Interface Standard in OGC)	RDA: Array Database Assessment Report (submitted and under adoption procedure)
number and market weight of implementations of an adopted standard (as far as can be determined)	OGC WCS (and, hence, CIS 1.0) known implementations (as before): rasdaman, GDAL, QGIS, OpenLayers, OPeNDAP, MapServer, GeoServer, NASA World-Wind, EOxServer, GMU; Pyxis, ERDAS, ArcGIS OGC WCPS known implementations: rasdaman, Netgis Server, GeOrbis
Reference Implementations contributed	None in this period
instructional material produced in support of standards (webinars, slide sets, ...)	Continuous maintenance o OGC standards wiki, http://myogc.org/go/coveragesDWG

Extra work not in DoA, but necessary in the course of standardization work	Several unforeseen work items in OGC and ISO, in particular extended discussions
--	--

4 Conclusion

EarthServer-2 has achieved a remarkable footprint in the field of “datacube” standardization, manifesting European impact in the Computer Science field which traditionally is dominated by the US.

Notably, all standards adopted have been implemented in rasdaman, demonstrating feasibility, and, at the same time, making rasdaman reference implementation of these standards wherever the standardization bodies support this notion ([OGC WCS](#), [OGC WCPS](#), [INSPIRE WCS](#) – see Figure 2). Deployment of large-scale services in WP4, totalling 3 PB, has convincingly shown that the coverage standards allow for scalable, up-to date implementations and services.

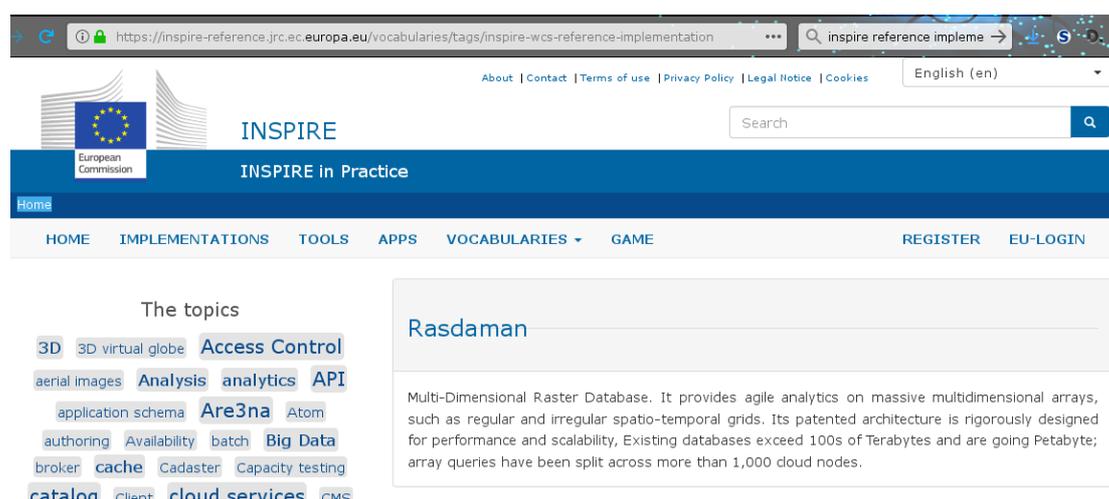


Figure 2: INSPIRE Reference Implementation, rasdaman

Specifically in the reporting period, datacubes are gaining substantial traction in the Earth science community, driven no longer only by EarthServer, but in particular through Australia and the US. Both have teamed up to establish Open Data Cube as a (python script based) tool³ which they market aggressively in international projects (Colombia, Vietnam, Uganda, and others) and with support they gained from organizations like CEOS. Although ESA currently leads CEOS unfortunately no effort has been undertaken to counter this domination by pointing out availability (and proven superiority, see ADA:WG report) European technology.

, these standards provide a solid framework for their interoperable implementation – and, in fact, a large and growing number of both open-source and proprietary tools [implement WCS](#): rasdaman, MapServer, EOxServer, GeoServer, GDAL, QGIS,

³ Alone in 2017 ODC managed to mobilize 20m\$ funding.

OpenLayers, OPeNDAP /Hyrax, George Mason University, NASA GSFC, Envitia ChartLink, Pyxis WorldView Studio, Constellation-sdi, Pixia, ESRI ArcGIS, etc.

Likewise SQL/MDA is being announced widely by the SQL group itself, and is seen as a game changer in future, as it allows a seamless combination of (relational) metadata and array data in Earth, Space, Life sciences and beyond⁴.

EarthServer-2 has substantially enabled the standardization work (as well as RDA work). The ambitious goals of EarthServer-2 ultimately have been achieved to a large extent (in face of the external impact factors of standardization, such as the rise of “datacubes” and the subsequent international competition on seizing the term). Several standards are in the adoption pipeline close to the end: ISO SQL/MDA, ISO 19123-2, OGC WCS 2.1. Further standards are in the pipeline and will be produced in continuation of the EarthServer-2 work, in particular ISO 19123-1.

Based on this achievement, work will be continued by JacobsU and RAS, supported by both national, ESA, and European funding sources found meantime.

⁴ In fact, a similar integration for XML and arrays (not being standardized for now) has been achieved by Athena Research / CITE during EarthServer-1 and -2.