ORIGINAL ARTICLE

Open Access

INSPIRE coverages: an analysis and some suggestions



Peter Baumann^{1*} and Jordi Escriu²

Abstract

In its quest for a common European Spatial Data Infrastructure INSPIRE has also addressed the category of spatiotemporally extended coverages, in particular: raster data. The INSPIRE definition of coverages is similar to the OGC and ISO standards, but not identical. This deviation from the common standards disallows using standard off-theshelf software. Further, as we find some technical peculiarities do not lead to the desired effects. In this contribution we compare INSPIRE coverages with OGC/ISO coverages, spot the differences, explain their consequences, and propose a minimal set of changes to the INSPIRE model achieving to harmonize it with OGC/ ISO and to remedy the issues found.

Keywords: INSPIRE, Coverages, Datacubes, WCS, OGC, Rasdaman

Motivation

Understanding and predicting natural processes form a vital source of insight for humanity at large – after all, the past, current, and future state of our planet determines human survival, welfare and prosperity. Not surprisingly, several global initiatives – including the United Nations Sustainable Development Goals [1], Sendai Framework for Disaster Risk Reduction [2], Paris Agreement on Climate Change [3] - have set out to enhance our insight on the Earth system processes and their impact on critical societal benefit areas, such as biodiversity, disaster resilience, food security, and water resources [4].

An important prerequisite for tackling the grand challenges as well as performing day-to-day management are Spatial Data Infrastructures (SDIs), with INSPIRE being one of the most widely known large-scale efforts. In the quest for a common, homogenized Spatial Data Infrastructure for Europe [5], the INSPIRE legal framework aims at providing standardized geo data interfaces to achieve interoperability of data and services across European institutions at all levels of administration. INSPIRE roadmap has prioritized addressing interoperability and provision of metadata and vector data, thoroughly used in thematic areas corresponding to Annex I of the

Large-scale deployments, such as with EarthServer [6], have proven coverage-based technologies can be extremely useful and powerful in thematic domains such as meteorology, orthoimagery, and elevation. On the other hand, however, many further thematic communities where potential for coverage data support is visible – such as geology and geo statistics – are still in an initial phase of understanding and using such technologies. Nevertheless, the ample presence of coverage-type data in the delivery formats selected by INSPIRE demonstrates the great expectation coverage technologies provoke across almost all thematic communities, having the challenge of Big Data as a background.

For a common modelling of spatio-temporal raster data the established concept of a *coverage* is utilized. Coverages are defined by OGC and ISO in the respective standards [7–12] as a unifying paradigm for spatio-temporal regular and irregular grids, point clouds, and general meshes. Based on the abstract, conceptual coverage model of ISO 19123:2005 [11, 12] (operating

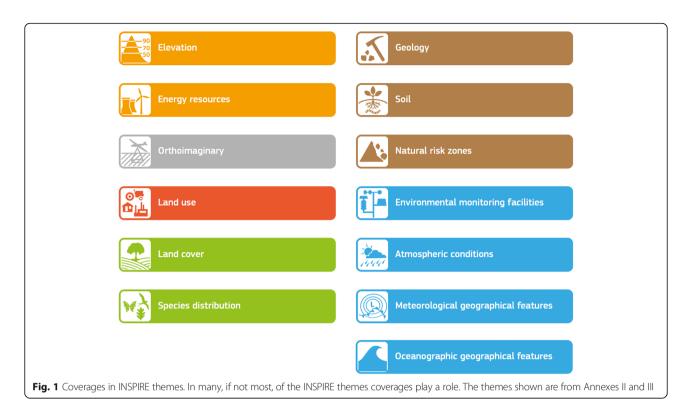
Full list of author information is available at the end of the article



Directive; the third basic geo data category, raster data, has been addressed only later in the roadmap, when coming to Annexes II and III of the Directive. Each of these Annexes I, II, and III addresses a set of so-called *themes*, i.e., technically or thematically motivated groupings of data structures. As Fig. 1 shows use of raster data — and, more generally, coverages, as introduced below — in IN-SPIRE themes is manifold.

^{*} Correspondence: p.baumann@jacobs-university.de

¹Large-Scale Scientific Information Systems, Jacobs University, Bremen, Germany



on conceptual level), OGC has established interoperable standards for a coverage data model [7, 8] and service model [9, 10] around 2012 (operating on implementation level). A main benefit and potential of coverage data and services is that they are ready to be processed, as an input for performing spatio-temporal analyses in combination with other data layers (other coverages).

However, INSPIRE drafting teams, when applying the OGC data model standards on coverages for defining its own coverage data model in 2013, introduced particular coverage structures which present some deviations, probably due to misinterpretation and lack of ample experience on these standards. On the other hand, OGC's service model, Web Coverage Service (WCS), was adopted unchanged in 2016.

Recently, the first concrete experiences implementing these INSPIRE coverages has started unveiling some issues.

In this contribution, we report on our investigation on commonalities, differences, and harmonization opportunities for OGC/ISO and INSPIRE coverages. We provide a general overview about OGC coverage data and service standardization, and compare this to the specification of INSPIRE coverages and the status of their implementation. In particular, the article identifies existing challenges in the INSPIRE coverage model which should be resolved in order to prevent implementation issues once European member states start deploying coverage services at scale. A minimal number of changes to INSPIRE coverages are proposed in

support a successful INSPIRE implementation – legally mandated to be in place by 2020 – for all the themes using coverage data.

The remainder of this article is organized as follows. In Section II we give an overview on the OGC/ISO coverage standards, followed by a brief presentation of the OGC coverages reference implementation, rasdaman, in Section III. In Section IV we discuss INSPIRE coverages and deviations from OGC/ISO coverages, followed by a status brief on the state of INSPIRE implementation of coverages. A minimal remedying set of changes to INSPIRE coverages is proposed in Section V. Section VI concludes the paper.

OGC/ISO Coverage Standards

In this Section we give a brief overview on the OGC coverage standards suite, which is fully harmonized with ISO. A one-stop shop of links to the authoritative standards, previews of candidate standards, webinars, tutorials, etc. can be found on [13].

Coverage Data model

Raster data of all kind, from 1D sensor timeseries over 2D raster images up to spatio-temporal datacubes, are captured by the unifying modelling paradigm of *coverages*. Actually, the term coverage - a subclass of feature (i.e., geographic object) - is broader as it describes spatio-temporal regular and irregular grids (i.e., multi-dimensional datacubes), point clouds, and general meshes. Abstract,

high-level definitions are laid down in OGC Abstract Topic 6 [14] which is identical to ISO 19123 [11] (currently being reworked into 19,123–1); concrete, interoperable definitions are given in the companion standard OGC Coverage Implementation Schema (CIS) [7] which is adopted by ISO as 19,123–2 [12].

On a side note, CIS 1.0 was formerly known as "GML 3.2.1 Application Schema – Coverages", nicknamed GMLCOV; as this name caused considerable confusion in the community, OGC in 2015 decided to rename the standard to Coverage Implementation Schema (CIS) – hence, GMLCOV 1.0 and CIS 1.0 are synonymous.

Coverage are modelled like a function, given by the coverage's domain set ("at what coordinates can I find values?") and its range set ("what are the values?"). Further, to capture the full semantics of the values, a range type is added; this range type definition is based on the SWE (Sensor Web Enablement) Common [15] concepts so that sensor data can be transformed into coverages without information loss, thereby enabling seamless service chains from upstream data acquisition (e.g., through OGC SOS) to downstream analysis-ready user services (such as OGC WMS, WCS, and WCPS). Finally, an optional metadata bucket is part of the

coverage which can carry any additional information that may be relevant.

Such coverages can be represented in a variety of shapes – including tilings, coordinate/value pair lists – and formats - such as GML, JSON, RDF, a variety of binary encodings, as well as "containers" with mixed encodings. Hence, tools can request coverages in their favourite format from a server.

Let us inspect an orthorectified image timeseries, i.e., an x/y/t image datacube. For our purpose we choose XML as a printable ASCII format (Fig. 2).

In the Domain Set we find a General Grid whose Coordinate Reference System (CRS) in the *srsName* attribute is composed from EPSG:4326 (WGS84) contributing axes *Lat* and *Long*, plus a time CRS contributing axis *AnsiDate*. As OGC has decided to have all identifiers as URLs – and this CRS string acts as an identifier – CIS uses URL techniques for referencing and composing the single CRS that a coverage has: references to, say, the EPSG database and AnsiDate as well as composition using the *crs-compound* CRS constructor supported by the OGC CRS resolver [16]. In addition to the projected space/time coordinates we find the underlying Cartesian grid definition which establishes a 3x3x3 array of data.

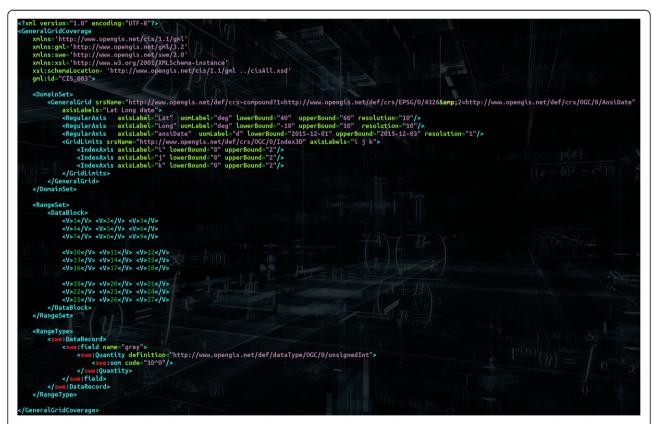


Fig. 2 Sample CIS 1.1 image datacube, in XML (source: OGC). Coverages can be encoded in various ways. Here an XML encoding is shown using the schema of OGC CIS 1.1 which is more "tidy" and comprehensible than older schemes. It clearly shows the mandatory components domain set, range set, and range type

These data we find in the Range Set's Data Block; as compared to earlier versions of CIS, values are not separated just by whitespace but enclosed in tags, thereby avoiding the specific micro syntax of introduced by GML 3.2.1 and its limitations (such as disallowing non-numeric date values). Hence, any XML parser can readily extract any value without extra coding effort.

Finally, the Range Type section shows that we have a "grey" (i.e., panchromatic) image type. Generally, the Rang Type carries rich information about the meaning of a pixel value, such as unit of measure, a link to a defining ontology, accuracy information, and more [15].

The Metadata container can hold any arbitrary data payload; see Fig. 3 for an example where the EO-WCS Application Profile standard makes use of the Metadata slot for attaching a contributing footprint polygon. The coverage does not know about its (application-specific) semantics, but it will duly transport it thereby guaranteeing data/metadata connection.

Actually, any number of such Metadata containers can be attached to a coverage, providing metadata on different facets. We will use this feature lateron in our proposal for harmonizing coverages and O&M.

Raster Coverage types

A core type of coverage for INSPIRE is the grid coverage family as it allows to model 1D timeseries, 2D imagery, 3D image timeseries, 4D weather data, and the like. The CIS 1.0 standard, with a tentative design decision to remain very close to GML 3.2.1, introduces three coverage types which we discuss in turn.

- GridCoverage. This is the historically first approach to gridded coverages. Today it is not recommended as there are limitations in the way coordinates can be expressed.
- RectifiedGridCoverage. This was the historically next approach, aiming at regular grids such as orthoimagery.
- ReferenceableGridCoverage. Originally, this was meant
 to capture "all other sorts of grids". Its definition,
 though, was left blank in GML 3.2.1; a corresponding
 change request adopted by OGC [17] has never been
 edited into the standard. Therefore, an extension
 specification to CIS 1.0 was established which provides
 an according definition for some irregular grid types.

In OGC Testbed-11 a comprehensive study on grid coverages was undertaken, resulting in a model generalizing the capabilities of GML 3.2.1 and 3.3, CIS 1.0, and SensorML 2.0 grid concepts in one comprehensive



Fig. 3 Sample CIS 1.1 coverage with metadata, in XML (source: OGC). This Figure shows an XML encoding of a coverage, following OGC CIS 1.1 as the previous one. However, in this case a metadata element has ben added providing further information about the contributing footprint. This extra information is defined in the EO-WCS standard

model, standardized as *GeneralGridCoverage* in CIS 1.1. Further, the CIS 1.1 General Grid Coverage definition allows cases not addressed by any of the other mentioned approaches, ultimately allowing the description of any multi-dimensional grid, with space, time, or any other semantics (such as spectral bands). The main difference is that it does not – more or less arbitrarily – differentiate *grid types* like the predecessor standards, but considers different *axis types* from which grids can be composed. For example, such grids can combine regular spatial axes (like an orthoimage) with an irregular time axis forming an image datacube.

Ultimately, axes forming grids need to describe the mapping between projected coordinates and the index coordinates of the underlying array representing the data grid. In CIS 1.1, the following axis types are distinguished:

- *Index axis.* Such an axis does not refer to a realworld situation, but resembles a simple Cartesian grid axis with integer coordinates the aforementioned mapping is 1:1.
- Regular axis. In this case, there is a simple linear dependency between projected coordinates and the underlying array coordinates. For the array mapping it is sufficient to store start and distance (i.e., resolution).
- *Irregular axis*. By abandoning the equidistance requirement we obtain an axis where we need to store all relevant grid coordinates in a list. Note that this still is efficient.
- Warped axes (distorted axis nests). Sometimes a grid
 is distorted in a way that each point needs to have
 its individual direct position stored explicitly. In
 such a situation, several axes (but not necessarily all)
 are engaged in the warped part of the grid. Note
 that storing the grid position of every grid point
 makes the domain set as large as the range set, or
 even larger.

Algorithmic transformation axes. In the most general
case, the coordinates of the points on a projected grid
are not any longer given by data, but by an algorithm.
Currently known are situations as described by
SensorML 2.0 which, for example, allows describing
mathematical models for estimating geolocations
from recorded camera data such as L0 swath imagery.
Typically, some ground control points are stored out
of which the transformation algorithm generates the
concrete grid coordinates. This gives complete
freedom in defining any sort of grid.

Recall that all these axis types can be freely combined in a GeneralGrid Coverage. Some examples are shown in Fig. 4. The CIS 1.1 standard comes with about 25 examples highlighting all main options the standard offers.

Coverage service model

While coverages can be served through a variety of interfaces - including WFS, WMS, and WPS - only Web Coverage Service (WCS) [9] offers comprehensive functionality, such as spatio-temporal subsetting and analytics. The modular WCS suite centres around a mandatory Core which defines subsetting, i.e., trimming (getting a cutout of the same dimension) and slicing (getting a cutout with reduced dimension) as shown in Fig. 5, as well as encoding in some user-chosen format. WCS Extensions add bespoke optional functionality, up to ad-hoc spatio-temporal analytics which we discuss below.

OGC provides a free, open compliance test suite which examines WCS implementations and the coverages delivered down to the level of single pixels, thereby ensuring interoperability across the large an increasing open-source and proprietary servers and clients [18].

Coverage analytics

Big Earth Data Analytics requires "shipping the code to the data" – which, however, begs the question: what kind of

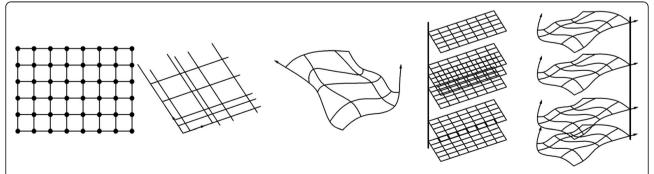


Fig. 4 Sample axis combinations in a General Grid Coverage (source: OGC). While older coverage standards attempt to classify grid types overall, OGC CIS 1.1 is more fine-grain and classifies axis types instead. Then, coverage grids can be built using different axis types in different dimensions, such as regular x/y axes and irregular time axis, resulting in an irregular ortho image timeseries datacube

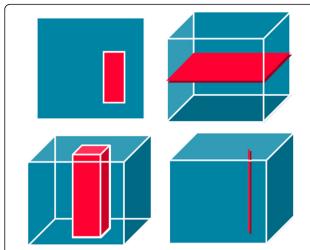


Fig. 5 WCS trimming (left) and slicing (right) (source: OGC). Specifically on datacubes (and point clouds) the WCS Core establishes simple (rectangular) subsetting, classified into trimming (left side – leaves dimension of the result unchanged over the object addressed) and slicing (right side – reduces dimension of the result). This WCS Core, therefore, defines the minimal WCS implementation. WCS extensions add further, optional functionality in a modular way

code? Certainly not procedural code (like C++ or python) which requires strong coding skills on the user side and also is highly insecure from a service operator perspective; rather, some safe, high-level query language is of advantage, in particular as it also enables strong server-side optimizations, including parallelization, distributed processing, and use of mixed CPU/GPU, to name a few. The ISO SQL database language acts as a shining example here.

Further, to provide data ready for spatial and temporal analytics, a semantically adequate integrated space/time coordinate handling is necessary. OGC already in 2008 has standardized the Web Coverage Processing Service (WCPS) geo datacube analytics language [10, 19] as part of the WCS suite. Today, WCPS is being used routinely in Petascale datacube services, e.g., in the intercontinental EarthServer initiative [6].

In a nutshell, the WCPS language [10] works as follows (see [19] for a comprehensive presentation). Adopting a syntax tentatively close to XQuery/XPath (so as to allow an integration with XML or JSON based metadata retrieval), the *for* clause defines iterations over coverage objects; the *where* clause allows filtering of coverage combinations to be retained, and in the *return* clause, processing (including coverage fusion) of the coverages is specified, including an indication of the delivery format of the result. A query like "From MODIS scenes MI, M2, M3: difference between red & nir, as TIFF, but only those where nir exceeds 127 somewhere" can be written as. for \$c in (M1, M2, M3) where some (\$c.nir > 127) return encode (\$c.red -\$c.nir.image/tiff")

Kakaletris et al. have implemented a coupling of this language based on a combination of the rasdaman array engine and an XPath processor, leading to seamlessly integrated data/metadata queries [20].

Coverage Data and service model interplay

Both the coverage data and the service model are established in a modular way: a *core* defines the mandatory components common to any implementation of the standard while *extensions* add optional functionality facets which an implementation may or may not support. In any case, extensions are clearly indicated, and compliance tests allow checking validity of coverage services down to the level of single pixels/voxels. The complete "Big Picture" of the coverage standards suite is shown in Fig. 6.

This modular data and service concept also defines mechanisms for standards-conformant extensions. Goal is that tools can decide to implement the (mandatory) core plus any set of extensions, while always remaining interoperable: any WCS implementation, for example, must support the basic subsetting mechanisms of the core, regardless of their choice of extensions. This requires some strict rules on extensibility. Essentially, the data and request structures must be retained unchanged by an implementation; the only extension point allowed is via the Metadata slot available with both CIS and WCS. For example, the WCS-EO Application Profile standard uses this extensibility mechanism to define satellite image coverages with specific add-ons, such as contributing footprints (Fig. 3). As we will see later one problem of INSPIRE coverages is that they modified the coverage definition by liberally modifying and extending data structures without respecting the extensibility rules of the standard.

Encoding of coverages is designed to support the various requirements different application scenarios mandate while retaining interoperability. One option is to use general-purpose encodings like XML or JSON. On the upside, such encodings are able to represent all coverage information, both generator and parser tools are readily available making them often convenient for use (such as using JSON when deriving 1-D diagram or 2-D image coverages for display in a Web browser); on the downside, such (typically ASCII-based formats tend to lead to rather voluminous representations, hence do not scale with large coverages. In such cases, efficient binary encodings like TIFF, NetCDF, or JPEG2000 provide an alternative – albeit at the cost of not retaining all information. Combining the best of both approaches - information completeness and volume efficiency – is possible through container formats which allow splitting coverages into sub-parts each of which gets encoded individually. For example, the coverage domain set, range type, metadata, etc. might get encoded in XML or JSON while the "pixel payload" - the

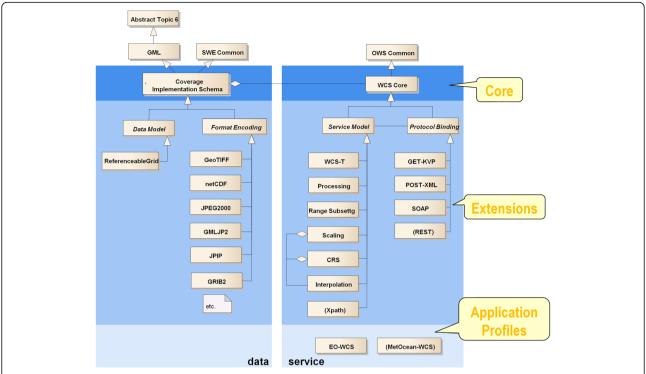


Fig. 6 Comprehensive synopsis of the OGC coverage / datacube standards suite (source: OGC). The "Big Picture" of the OGC coverage standards suite explains the separation into a data and service model, as well as the positioning of the individual format encoding, service functionality, and protocol binding extensions

range set – gets encoded in NetCDF. Suitable as a container format are formats like multipart MIME (such as used for email attachments), zip, GMLJP2, etc.

So far we have assumed that coverages are stored along their conceptual definition, with separate components for domain and range set. Due to manifold stakeholder requests this has been extended in version 1.1 to allow a fine-grain partitioning of coverages. Instead of the domain/range representation, a coverage may be recursively composed of sub-coverages. A second option, tuned towards timeseries generation, consists of organizing a coverage into a sequence of position/value pairs.

Note though, that both features should not be used naively as a storage organization in a server; it is a strength of the coverage model that it allows to organize data efficiently towards particular service functionality and access patterns while retaining all information and semantics. For example, timeseries analysis on an x/y/t coverage may suffer from inadequate performance if organized into horizontal slices (as done traditionally). Servers supporting adaptive tiling can be optimized towards any particular access pattern and, hence, are known to convey substantially better performance [21, 22].

This finalizes our brief inspection of the OGC/ISO coverage data model. With its information model synchronized with the sensor standards, coverages pave the way for orchestrating services advantageously. Of particular interest

may be the combination of SOS as an upstream data capturing and integration service which homogenizes all kinds of sensors. Through coverages, many such data can be integrated to larger, more user-centric units, like datacubes, and served through downstream WCS and all the clients supporting WCS already (Fig. 7).

Coverage Standards implementation

A broad investigation on datacube tools, including Array Database Systems and MapReduce-based systems, has been conducted by the Research Data Alliance (RDA) [22]. We choose to present rasdaman ("raster data manager") because it is official OGC reference implementation for coverages, the most comprehensive implementation of the CIS/WCS suite, and stands out in that it offers a high-level datacube analytics language which significantly eases building up streamlined services. Further, rasdaman has several scalability mechanisms built in to address the Big Earth Data challenge.

The rasdaman Datacube analytics engine

The rasdaman ("raster data manager") array engine is implemented completely from scratch, with every component hand-optimized to support management and retrieval on massive multi-dimensional arrays in a domain agnostic way. Its raster query language, *rasql*, leans itself towards SQL adding in array operators; actually, modulo

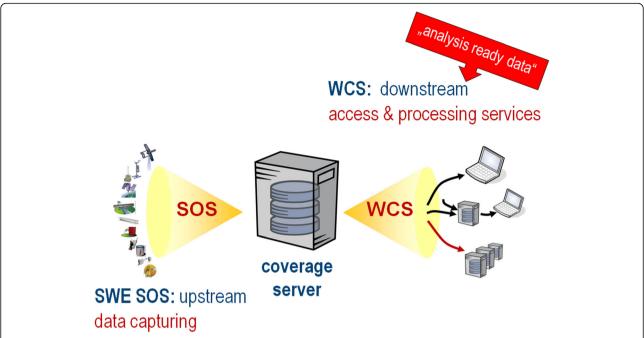


Fig. 7 Upstream SOS and downstream WCS service integration, with WCS providing Analysis Ready Data (source: OGC). Data can tream seamlessly from sensors, as captured upstream through SOS and encoded following O&M, into downstream coverage services where users see consolidated, analysis-ready data through WCS

syntax changes this language has been adopted by ISO as the SQL datacube extension [23].

On server side, query evaluation involves a variety of optimizations, including: adaptive data partitioning and compression, heuristic query rewriting, multi-core parallelization and mixed CPU/GPU processing, semantic caching, cost-based optimization, distributed processing (Fig. 10), etc.

Queries have been successfully distributed across more than 1000 Amazon cloud nodes [24]. By using the same principle across data centres, federations can be established. Figure 10 shows a visualization of actual federated query processing between the European Centre for Medium-Range Weather Forecast (ECMWF) in the UK and National Computational Infrastructure (NCI) in Australia - both running rasdaman - for determining heavy rainfall risk areas from precipitation data at ECMWF and the Landsat8 datacube at NCI.

Service implementation

The overall rasdaman system architecture is shown in Fig. 9. The core functionality of the framework developed consists of the rasdaman Array Database System for storage of remote sensing data and OGC WCPS interface standard for querying them. Rasdaman was selected as the core system of our implementation due to its proven robustness, novelty and efficiency in handling big imaging data [25, 26].

The developed framework in its current version provides geospatial services for precision agriculture, water quality monitoring and land cover mapping. For the provision of the services, the Landsat 8 and Sentinel 2 datacubes are utilized along with OGC WCPS interface standard and several implemented applications (external programs) in order to derive remote sensing information and produce respective colour maps that hold this information.

The overall system architecture comprises of a "data acquisition and pre-processing" stage for ingestion; the rasdaman database that hosts the EO data and supports OGC WCPS interface standard for processing and querying the image datacubes, a layer of server-side applications that produce the value-added maps and a Web Client that forms the front-end of the framework which handles interaction with the user. The system components, shown in Fig. 9, are detailed in the following.

For storage, arrays get automatically partitioned into tiles (Fig. 8) which – as opposed to, e.g., PostGIS Raster [22] – are not visible to the user, rather they are maintained automatically by the system. Administrators, though, can impact the tiling strategies and, hence, tune the system to the various workloads.

Figure 9 shows the overall system architecture of rasdaman. The core engine, represented by a set of parallelized worker processes called *rasserver*, is domain agnostic and supports managing and querying any sort of array

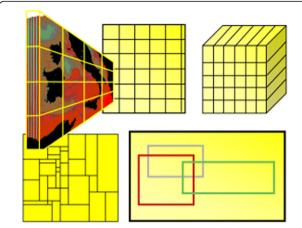


Fig. 8 rasdaman datacube partitioning examples, under administrator control via the rasdaman storage layout language. Datacubes need to be partitioned for storage (but for simplicity of service this partitioning should not be visible to users). As the currently only tool rasdaman allows arbitrary partitioning (here called tiling), thereby allowing to adjust to the users' access patterns. In the extreme case, only some areas of interests are indicated, and the system will automatically find an optimal tiling

datacubes, including life science data, cosmological simulations, and Planetary science [27], among others. The engine supports both multi-Core/GPU parallelization and distributed processing, effectively allowing to establish federations between data centers and even satellites [28] without a single point of failure (Fig. 10); a sample federation between the European Centre for Medium-Range Weather Forecast (ECMWF in Reading/UK and National Computational Infrastructure (NCI) Australia in Canberra have been demonstrated

live (Fig. 11). Queries have been split massively across multiple compute cores [24].

Data can be stored in a conventional database as BLOBs, such as in PostgreSQL (avoiding the overhead and particularities of PostGIS Raster, though). Alternatively, an internal storage format can utilize the file system directly, leading to some performance increase of about 2x. For massive Petascale data sets where data import (i.e., copying) is not an option data can remain in their existing archive, and by way of configuration rasdaman can adjust to any pre-existing archive structure; in this case, processing is done in-place without data copying, with no difference for the user at the query interface.

Geo semantics is provided through a dedicated layer which knows about space and time coordinates, regular and irregular grids, etc. Functionality of this layer is provided through the OGC interface standards WMS, WCS, WCPS, and WPS.

Current state of INSPIRE COVERAGES

Use of coverages in INSPIRE

Coverages are widely used in different thematic domains in the scope of the INSPIRE Directive (see Fig. 12):

- Themes from INSPIRE Annex II: Elevation (EL), Land cover (LC), Orthoimagery (OI), Geology (GE);
- Themes from INSPIRE Annex III: Soil (SO), Land use (LU), Natural risk zones (NZ), Environmental Monitoring Facilities (EF), Atmospheric conditions (AC), Meteorological geographical features (MF), Oceanographic geographical features (OF), Energy resources (ER), Species Distribution (SD);

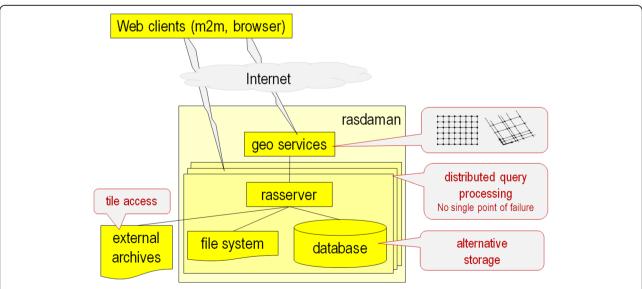


Fig. 9 rasdaman overall architecture. Being a full-stack implementation of an Array DBMS rasdaman is hand-crafted for multi-dimensional arrays. Tiles can be stored in various ways, including remaining "in situ" in some existing archive without copying

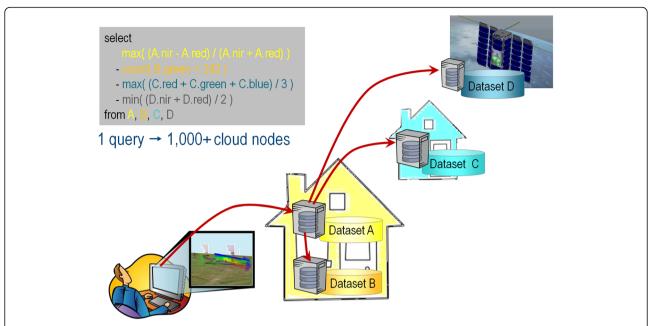


Fig. 10 rasdaman transparent distributed query processing. In a rasdaman federation queries get automatically split to the various servers involved, and likewise intermediate results get aggregated automatically. Hence, users see only the result and are not bothered with any of the processing steps involved in data fusion

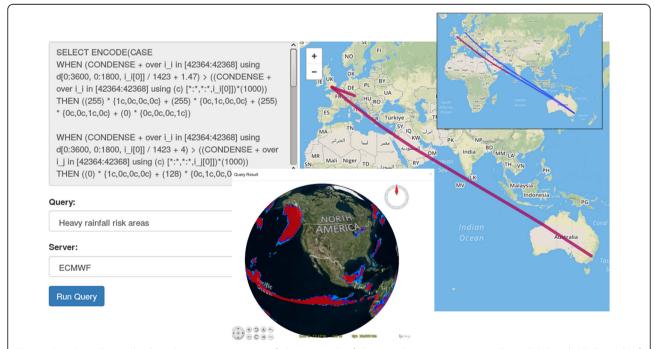


Fig. 11 Sample rasdaman distributed query processing in a federation. In this federation demonstration executed in WCMAF and NCI the path of a query sent from Germany to ECMWF/UK is drawn which forks off a subquery to NCI Australia, returning a precipitation/Landsat8 fusion product, with the result shown center bottom); top right: same query sent to NCI which splits a subquery to ECMWF, returning the same result



Fig. 12 INSPIRE Themes for which coverages data shall or may be provided. This graphics shows the data set candidates for being modelled as coverages in INSPIRE. See article text for details

The Statistical Units (SD) and Habitats and biotopes themes also have discussed the possibility of using coverages for data provision, but finally discarded it due to the immaturity of the corresponding standards at the time of drafting the INSPIRE technical guidelines in 2012.

Coverages play a role in all these themes, either as a requirement (as per the provisions of the Implementing Rule on spatial data sets and services [5]) or as an option for data provision. Uniformly, coverages shall be provided according OGC CIS 1.0 [8].

Coverage Data modelling and content

The INSPIRE technical guidelines of the different themes were drafted as conceptual data product specifications. Therefore, the corresponding data models - being part of them - share the same level of abstraction: they are abstract and conceptual. For this reason, in addition it was necessary to specify concrete rules for the delivery and encoding of data, either as part of the own data specifications or in separate INSPIRE guidance documents (e.g. INSPIRE D2.7 – Guidelines for the encoding of spatial data v3.3 [29]).

Despite this approach was also applied to the INSPIRE application schemas which serve to model coverage data across different themes, the rules specified were not complete enough to assure the interoperability of data (coverages) and related services - due to a possible

misinterpretation and lack of ample experience on the standards for implementing coverages. In order to avoid disparate modelling results and to share a common abstract structure at INSPIRE level, the *INSPIRE Generic Conceptual Model v3.4* (INSPIRE GCM - D2.5 guideline document [30]) recommends using the *Base Model for Coverages* published as part of the *INSPIRE Data Specifications – Base Models – Coverage Types 1.0* (D2.10.2 guidance document [31]). This so-called *seed model*, aimed at providing cross-theme harmonization, is composed of common classes and application schemas for coverage types aimed to be reused by those INSPIRE thematic data models using coverage data. It was developed in accordance with the ISO 19123:2005 [11] standard, i.e. at abstract level.

The main abstract *Coverage* feature type defined in these common schemas carries the basic properties of a coverage to be implemented according OGC CIS 1.0: domain set, range set, range type and metadata bucket. The *Coverage* feature type is further specialized in two abstract classes: *RectifiedGridCoverage* representing regular grids, and *ReferenceableGridCoverage* for irregular grids.

However, despite the original purpose of the mentioned seed model, modellers in the various Thematic Working Groups still came up with heterogeneous ways of deriving their coverage data models from it. As a first observation, this was due to the high level of abstraction of the seed model, based on ISO 19123, as explained earlier.

As an illustration, Fig. 13 shows the various different ways used at the moment in the INSPIRE application schemas to define thematic gridded coverage data (see the set of INSPIRE technical guidelines for the themes, [32–45]), as described below:

- As regular grids, specializing a coverage data class from RectifiedGridCoverage class of the Base Model for Coverages, the seed model onwards. This is the case for Elevat-ion (EL), Land cover (LC), Orthoimagery (OI), Soil (SO), and Land use (LU).
- As regular or irregular grids, by reusing either RectifiedGridCoverage or ReferenceableGridCoverage classes of the seed model, as data types. This is done in Geology (GE) for the description of the hydrogeological surfaces (piezometric state).
- As regular or irregular grids, specializing four different coverage data classes from

- CoverageByDomainAndRange class of the seed model, and constraining the domain of the coverages to CV_RectifiedGrid or CV_ReferenceableGrid classes defined in ISO 19123 [11], respectively. This approach has been taken in Natural risk zones (NZ).
- As regular grids, specializing a coverage data class from *CoverageByDomainAndRange* class of the seed model, and constraining its domain to *CV_RectifiedGrid* class (ISO 19123 [11]). This has been done in Energy resources (ER) and Species distribution (SD). The application schema in the latter theme, SD, is currently deprecated. It also allows using multi-point (*GM_Multi-Point*) and multi-surface (*GM_MultiSurface*) discrete coverages.
- As regular or irregular grids provided as discrete observation coverages, i.e. gridded data specialized observation types applying the ISO 19156:2011

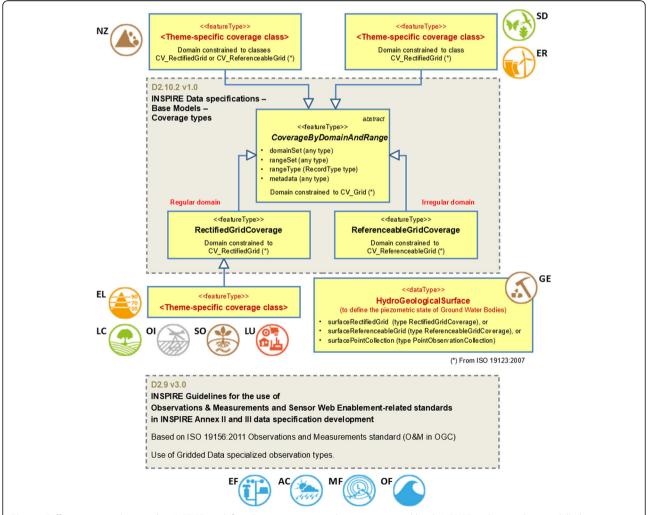


Fig. 13 Different approaches used in INSPIRE to define thematic coverage data at conceptual level. INSPIRE in the past has modelled coverages with some handwaving and sometimes deviation from the OGC principles, with specialization used for deriving various coverage subtypes for the individual themes

Observations and Measurements standard (O&M in OGC), following the INSPIRE Guidelines for the use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE Annex II and III data specification development (D2.9 v3.0 guidance document [46]), as required by the INSPIRE GCM. This is the case with Environmental monitoring facilities (EF), Atmospheric conditions (AC), Meteorological geographic features (MF) and Oceanographic geographic features (OF).

In practice, a number of INSPIRE-defined properties were identified by the Thematic Working Groups (TWGs) which needed to be added in the context of the European spatial data infrastructure. They were appended to each of the theme-specific coverage feature types represented in Fig. 13, adding contents not foreseen in the OGC CIS 1.0 implementation standard. From now on we refer to such additional contents as *INSPIRE extensions*.

As an illustration of such INSPIRE extensions we inspect the application schemas of Elevation and Orthoimagery themes. Most of the extensions we find there correspond to properties which were identified and considered useful during the INSPIRE drafting process to manage coverage data. They are not reflected in OGC CIS 1.0, but INSPIRE may submit change requests to OGC in the future in order to take these items into account in eventual revisions of CIS. One good example is a geometry/value pair list representation of coverages which is particularly useful for 1D timeseries. Meantime, this has been included in CIS 1.1 independently, plus further options like tiling. In other cases, INSPIRE introduced information elements at conceptual level without specifying their mapping to the corresponding elements already in the OGC CIS 1.0 [5] (e.g., the relationship of inspireId to the coverage identifier) and the SWE Common [15] implementation standards.

Each of the INSPIRE conceptual data model elements included in Table 1 are discussed in turn below. The *inspireld* attribute is the INSPIRE unique and persistent external identifier of the coverage. It might be mapped to the *gml:id* attribute which each GML feature has, and so does CIS. Note, however, that XML (and, hence, GML) only mandate uniqueness of an id attribute within a given coverage document. However, it remains unclear what the *inspireld* value should be in case the coverage is obtained as a subset from some coverage which has its own *inspireld*. In general, it may not say anything about service or planet wide uniqueness. Therefore, the rules governing this mapping need to be analysed further.

Additionally, however, since this item has been identified as required (mandatory) in the case of the EL and OI themes, the coverage also explicitly carries an

inspireId attribute – see Requirement 61 and Recommendation 33 of the INSPIRE Generic Conceptual Model (D2.5) v3.4 [30]. Hence, *inspireId* attribute may be also considered as an INSPIRE extension.

The *beginLifespanVersion* and *endLifespanVersion* attributes represent the lifecycle information reused by all INSPIRE themes; they are INSPIRE specific and do not convey any correspondence to GML or CIS constructs.

The *domainExtent* attribute serves to describe the approximate spatio-temporal extent of the coverage. It can be mapped to the *gml:boundedBy* property inherited by CIS coverages from *AbstractFeature* by either implementing *gml:Envelope* or *gml:EnvelopeWithTimePeriod*.

Since *inspireId*, *beginLifespanVersion*, *endLifespanVersion* and *domainExtent* attributes of INSPIRE coverages constitute extensions to CIS, while being required for EL and OI themes, they can be included within a metadata instance provided in the metadata bucket of the coverage.

An INSPIRE EL or OI coverage is provided with an aggregation relationship (contributingElevationGridCoverage in EL; contributingOrthoimageCoverage in OI), for referencing other coverages (zero or more) which, together with the former one, represent a set of sub-coverages being virtually aggregated to form one super-coverage; such constructs are referred to as aggregated or nested coverages. Since the super-coverage does not exist in itself its spatial extent can only be derived by composing the footprints of the associated sub-coverages. Such footprints can be disparate, heterogeneous, disjoint, and even overlapping. For this reason, a contributing-Footprint attribute is foreseen which delineates the areas (set of pixels) in each sub-coverage which effectively contribute to the virtual super-coverage. Technically, such an attribute adopts the GM_MultiSurface type which allows providing a multi-polygon area (the area of the sub-coverage visible from a super-coverage perspective). The attribute is included within the corresponding association class in the EL and OI data (ElevationGridCoverageAggregation models OrthoimageAggregation in OI).

In passing we note that the virtual super-coverage grid cell values (pixels) are those within the area composed by the aggregation of each sub-coverage's *contributingFootprint*.

Interestingly, talks to stakeholders have revealed that only the European Space Agency (ESA) requested this coverage aggregation mechanism to be considered in INSPIRE. On the other hand, complaints about undue complexity of the model have been reported manifold.

The *mosaicElement* association relationship in the OI theme serves to provide information describing the orthoimage mosaic, composed of a contiguous set of

Table 1 Mapping EL&OI INSPIRE coverage to OGC/ISO CIS 1.0 elements

INSPIRE Conceptual data model	Mapping to OGC CIS v1.0
Coverage element: type [cardinality]	
< <stereotype>></stereotype>	CIS (gmlcov:) / GML element (gml:)
Element type	
ElevationGridCoverage (type RectifiedGridCoverage as defined in D2.10.2 [18]) < <featuretype>> Spatial object OrthoimageCoverage (type RectifiedGridCoverage as defined in D2.10.2 [18])</featuretype>	gmlcov:RectifiedGridCoverage
<featuretype>> Spatial object domainSet: Any [1]</featuretype>	
(type constrained to CV_RectifiedGrid) Attribute	gmlcov:RectifiedGridCoverage.domainSet [1] (type gml:RectifiedGrid)
rangeSet: Any [0*] (type duly constrained in each theme; it represents each of the coverage values) Attribute	gmlcov:RectifiedGridCoverage.rangeSet [1] (type gml:RangeSet; represents the set of coverage values)
rangeType: RecordType [1] Attribute	gmlcov:RectifiedGridCoverage.rangeType [1] (type swe:DataRecord)
metadata: Any [0*] Attribute	gmlcov:Coverage.metadata: Any [0*]
inspireld: Identifier [1] Attribute	gml:id property (To be further analysed)
domainExtent: EX_Extent [1*] Attribute	gml:boundedBy / gml:Envelope / gml:EnvelopeWithTimePeriod
beginLifespanVersion: DateTime [1] < <voidable>> Attribute</voidable>	Not matched (INSPIRE Extension).
endLifespanVersion: DateTime [1] [01] (Cardinality [1] in EL; [01] in OI) < <voidable>> Attribute</voidable>	Not matched (INSPIRE Extension).
contributingElevationGridCoverage [0*]	N
Aggregation relationship contributingOrthoimageCoverage [0*] Aggregation relationship	Not matched (INSPIRE Extension).
ElevationGridCoverageAggregation Association class	Not matched
OrthoimageAggregation Association class	(INSPIRE Extension).
propertyType: ElevationPropertyTypeValue [1] Attribute	Not matched (INSPIRE Extension).
surfaceType: SurfaceTypeValue [1] Attribute interpolationType: InterpolationMethod [1]	Not matched (INSPIRE Extension). Not matched
Attribute footprint: GM MultiSurface [1]	(INSPIRE Extension).
<pre><<voidable>> Attribute</voidable></pre>	Not matched (INSPIRE Extension).
name: CharacterString [01] < <voidable>> Attribute</voidable>	Not matched (INSPIRE Extension).
phenomenonTime: TM_Period [01] < <voidable>> Attribute</voidable>	Not matched (INSPIRE Extension).
mosaicElement [0*] < <voidable>> Association relationship</voidable>	Not matched (INSPIRE Extension).

Legend:



possibly irregular image areas (delimited by seamlines) whose range sets come from different orthorectified scenes, all having the same grid and the same range type. Each of these scenes is the result of orthorectifying an image (which is often in a conic projection) captured by means of an airborne sensor, at a specific date of acquisition. Such source images partially overlap each other, providing, for example, stereoscopic capabilities in the overlapping areas between two contiguous scenes. Therefore, the orthoimage coverage constitutes a mosaic composed of selected areas from the orthorectified scenes used for producing it.

Each of the elements in the orthoimage mosaic is represented by the *MosaicElement* feature type foreseen in the INSPIRE OI data model, and linked from the mosaic through the *mosaicElement* association relationship. Such feature type allows providing the area (*geometry* attribute) and the acquisition date related to each orthorectified scene participating in the mosaic (*phenomenonTime* attribute).

This information is rather technical and rarely relevant for users. Consequently, it does not normally form part of the final coverage product.

None of the concepts provided in INSPIRE EL and OI for describing coverage aggregation, and orthoimage mosaic information are foreseen in CIS 1.0, and therefore also constitutes extensions to the standard. While CIS 1.1 does support nested coverages, it places strict constraints for maintaining homogeneity of the results, i.e.: allowing users to see the super-coverage as one single, coherent piece of information without extra complexity.

Apart from the elements described above, INSPIRE themes also add some theme-specific attributes to their corresponding coverage subtypes derived from the seed model. They may serve for different purposes, such as searching, filtering and retrieving coverage data across the pan-European spatial data infrastructure. For our two themes considered in this paper, OI and EL, we find the following:

- In the EL coverage data model, the propertyType
 attribute identifies the type of elevation property
 described by the coverage, i.e. height or depth;
 surfaceType attribute identifies the type of surface
 represented, i.e. either a Digital terrain Model
 (DTM) or a Digital Surface Model (DSM).
- In the OI coverage data model, the *name* attribute is a free text name which describes the orthoimage coverage; *interpolationType* attribute specifies an interpolation method meaningfully applicable to the coverage; *footprint* attribute defines the refined extent of the coverage, which delineates areas of the coverage with no null values; further, *phenomenonTime* attribute is a description of the observation/acquisition temporal extent of the coverage.

All these theme-specific attributes neither map to any existing GML nor CIS properties. Therefore, they constitute INSPIRE extensions as well which can be provided within the coverage metadata bucket, as described earlier.

Coverage Data encoding

Encoding for coverage data in INSPIRE is based on CIS 1.0 which establishes both an XML encoding, binary encodings, as well as combined representations.

In the case of specific thematic domains, INSPIRE allows other alternative formats. Examples include the BAG format used to provide bathymetry values within the marine community. Adding further coverage encodings is fully in line with the OGC CIS philosophy and perfectly consistent as long as the mapping of the format elements (like TIFF tags) to the corresponding coverage constituents is defined clearly. INSPIRE actually might take initiative to establish coverage format extension standards in OGC as has happened in the past, for example, with the GRIB2 format needed by the weather and climate modellers and meantime adopted by OGC as one additional coverage encoding.

Coverage services

INSPIRE allows delivering coverage data through various Web service APIs. Basically, two alternative approaches are foreseen, download of predefined datasets and WCS. We inspect both in turn.

Predefined datasets present coverage data split into pieces based on existing distribution units (like map sheets). These distribution units are then delivered using the same download services as INSPIRE vector data (Atom or WFS). In such cases, the distribution units are equivalent to a specific tiling scheme - in other words, pixel data are accessible only in this predefined packaging, without any processing such as spatial or temporal subsetting, band extraction, scaling, CRS transformation, etc. When handling high volumes of coverage data, splitting it into digestible pieces or tiles is the only way to achieve efficiency using Atom or WFS services as these service paradigms allow only delivery of complete coverage objects, but do not make use of subsetting. The consequence is that providers prefer to offer many small coverages at a granularity they consider adequate, and users see a myriad of files instead of a single consolidated object. Choice of splitting, such as tiling schema and parameters is up to each data provider, and consumers must be informed about it for determining the right objects, i.e.: tiles in this case.

Although this approach is often used by data providers less experienced dealing with coverage data use of predefined datasets is not the most flexible and powerful way for delivering it. Such kind of interfaces represents simplistic implementation techniques which are geared towards simple delivery of preprocessed files, while being inflexible for certain user demands. State of the art, however, is to offer seamless maps and allow arbitrary cut-outs and zoom levels to be retrieved in one go.

This is what the WCS service model brings in, with its subsetting and reformatting capabilities in the core, plus a set of optional extensions for various functionality facets. As this OGC standard is specifically designed for coverage data these may be offered homogenized as a single coverage based on a common grid (i.e., common space allocation defined in a specific CRS, with predefined resolution levels) and common semantics (i.e., the same range type). Such coverages appear as single logical objects to the user although they have possibly been composed from a large number of input files. Users have the opportunity to request and exploit any desired subset, regardless of any tiling schema the provider may choose - meaning that existing archive structures can remain unchanged and partitioning can be optimized towards particular workloads (such as preferred spatial or temporal access). In any case, this complexity remains hidden to the service user. Results can even be tiled again prior to shipping when using some suitable data format such as TIFF which supports internal tiling.

INSPIRE WCS adopts OGC WCS 2.0 without modifications. Any INSPIRE WCS must implement WCS Core (as OGC requires), plus the WCS Range Subsetting Extension. Optionally, an INSPIRE WCS may offer the OGC WCS CRS Extension and the OGC WCS Processing Extension which provides a fully-fledged spatio-temporal geo datacube analytics language.

Access control

For efficiency reasons, it is highly advisable to limit the maximum volume of data that may be requested to the service in a single query. Generally, various relevant server capacity parameters can be fine-tuned, such as transmission bandwidth, but possibly also client-side constraints such as RAM and disk space available for holding downloads.

For vector data and metadata this is not an issue normally. However, given the Big Data that coverages typically represent this is a new requirement that service providers may want to see addressed. In fact, the field of access control, quota, and billing has not been sufficiently addressed by any of the geo service standardization bodies; vendors have to provide bespoke solutions for now. Recognizing this gap OGC is conducting first conceptualization and experiments in the 2018 Testbed-14 [47].

Summary of INSPIRE Coverage issues

In summary, our interoperability investigation has led to the following observations. To emphasize again, we do not judge on the feasibility of modelling contents, such as the rationale for specific attributes in INSPIRE; we accept the information needs of INSPIRE coverages as a given. Instead, we exclusively focus on harmonization issues, and here we have found the following inconsistencies of INSPIRE coverages with OGC/ISO coverage standards:

- Need to analyse the mapping between the *inspireId* attribute and the *coverageId*. While the *coverageId* attribute in OGC/ISO is only defined locally for identification within a service, *inspireId* has a wider scope a unique and persistent identifier across the European infrastructure. This requires establishing specific rules governing their formal relation.
- INSPIRE coverage definition is heavily based on structural extension mechanisms not compatible with OGC/ISO coverages (where extensibility concepts are clearly defined to ensure interoperability in presence of modular flexibility). We found the following problematic patterns preventing interoperability:
- Added elements (in the XML sense). For example, inspireId, beginLifespanVersion, endLifespanVersion, and nilReason have been added directly under the document root. OGC CIS compliant implementations will not validate such coverages, or if no individual instance validation is performed, which is common practice for performance reasons they will likely ignore these elements.
- Some elements have been introduced in the INSPIRE coverage models which conceptually maps to existing OGC CIS elements. For example, the *domainExtent* definition of INSPIRE overlaps the e.g. *boundedBy* and used in CIS. Consequently, a complete list of encoding rules including the mapping between the INSPIRE coverage elements and the CIS coverage implementation elements need to be analysed and stated in INSPIRE technical guidelines.

Additionally, the liberal use of subclassing in Fig. 13 may generate, at implementation level, structures incompatible with OGC/ISO coverages. This makes vanilla WCS implementations likely to break, both client and server. In particular, while a server can choose to implement optional features on a case-by-case basis, clients need to be prepared for any combination of optional parameters — hence, any additional optional parameter complicates implementation significantly.

• Concepts which are used by INSPIRE coverages from different themes are modelled independently in their conceptual data models (i.e. there is a lack of cross-theme harmonization). For example, *contributingElevationGridCoverage* and

contributingOrthoimageCoverage resemble the same mechanism and, therefore, should be defined centrally so as to make this concept available to other Annex II and III themes likewise; as it stands, they in turn need to define parallel concepts. However, this per se is not an interoperability issue and, hence, does not require attention from this perspective.

 Conceptual extensions over the OGC/ISO coverage model. For example, aggregated (also called nested) coverages complicate the service model substantially while only required by some high-end applications and users, like for the European Space Agency.

These deviations altogether may cause several adverse effects:

- Lack of interoperability: in the absence of a complete list of encoding rules, mapping the conceptual elements in INSPIRE coverages classes to the corresponding CIS 10 elements, data providers may end up using different and incompatible structures due to diverse interpretations of the data specifications; this imposes a harmonization load to client tools and likely even requires educated human intervention on user side.
- Unexpected issues may arise when using standard
 off-the-shelf WCS servers and clients as outlined,
 tools (in particular: clients) may break encountering
 structures which are not expected. Still, this is better
 than the next possible effect because the problem
 will be signalled.
- OGC compliant implementations may silently ignore certain coverage parts. This is the most dangerous situation because wrong results can be generated and go unnoticed as both service providers and clients expect tools to work, based on OGC compliance tests passed.

Altogether, the INSPIRE approach obviously has been guided by a strong GML perspective. Now this may be suitable for vector information any information elements may be modelled and encoded in GML by deriving from the corresponding XML specific schemas. However, implementations of coverage standards will not solely reside in GML world, for two reasons: first, GML is but one representation format, and by far not the prevailing one, say, for raster images; therefore, the very specific methods of GML type derivation will not carry over to the general coverage implementation.

Second, an implementation will use the encoding – such as GML – merely as a frontend. Modifications in the schema, as done by INSPIRE, will not automatically change the internal algorithms and data structures of a WCS implementation – all the server-internal

processing (such as subsetting, band extraction, CRS transformation, analytics) will be done by code which has no idea about the GML schema. Hence, a GML-centric view will not match with existing image processing tools such as, for example, GDAL. Acknowledging the long heritage of image processing tools OGC CIS has established a derivation and subtyping method which ensures consistency across WCS Core and Extensions, for example, and is compatible with existing tools as the variety of WCS implementations shows, such as rasdaman, MapServer, GeoServer, EOxServer, GDAL, QGIS, ESRI ArcGIS, and more [48].

Summing up, as far as coverage data is concerned, any items of information modelled and not foreseen in OGC CIS 1.0 may lead to different implementations by data providers which may also prevent correct functionality of WCS services. We will propose INSPIRE adjustments in the next section.

Proposed changes to INSPIRE coverages

In this Section we propose a small set of minimally invasive changes to the INSPIRE coverage model in order to remedy the shortcomings stated, and to establish coherence, down to the level of compliance testing, with OGC/ISO coverages.

Some of these issues are of semantic nature (such as *inspireId* handling, others are of syntactic nature (such as structural extension mechanisms incompatible from an OGC/ISO perspective). The former class can be handled by adding clarifications without changing existing INSPIRE regulations. The latter class, though, requires changes. We will inspect these all in turn and propose remedies for each issue found.

However, in order to take profit of all the benefits and potential of data cubes (coverage data) and WCS, inter-operability shall be improved in the near future.

A number of suggestions are enumerated below.

Proper metadata placement

As discussed, INSPIRE has chosen to add coverage components in places where CIS disallows, thereby maintaining interoperability of CIS and WCS across all their extensions. To reap this benefit also for INSPIRE we suggest to shift all metadata which do not have a correspondence in CIS into the CIS metadata bucket. This metadata bucket should be designed as an embracing container – in GML it might be represented by an element *InspireMetadata* where all items are lined up. From the above investigation, examples include *inspireId* (see below for further discussion), *beginLifespanVersion*, and *endLifespanVersion*. A special case is *interpolationMethod*: in CIS 1.0 it needs to become a metadata element as well while in CIS 1.1 interpolation information is foreseen.

Encoded in CIS 1.1 XML such an INSPIRE metadata bucket would have the following structure:

```
<GeneralGridCoverage ...>
 <DomainSet>...</DomainSet>
 <RangeSet>...</RangeSet>
 <RangeType>...</rangeType>
 <Metadata>
    <InspireMetadata>
      <inspireId>...</inspireId>
      <br/>
<br/>
deginLifespanVersion>
        2018-07-20
      </beginLifespanVersion>
      <endLifespanVersion>
        2018-07-21
      </endLifespanVersion>
    </InspireMetadata>
 </Metadata>
</GeneralGridCoverage>
```

At the same time, this approach allows for further specialization (i.e., addition of extra elements) for the INSPIRE themes – the *InspireMetadata* container is under full, exclusive control by INSPIRE.

Placement of inspireld

Assuming that *coverageId* and *inspireId* play the same role they could simply be identified. The *inspireId* in a coverage shall be represented by its *coverageId*. However, we choose another approach as there are currently open questions, and it seems better engineering to keep both disentangled: we rather propose to add an *inspireId* item to the coverage metadata slot. First, the INSPIRE directive mandates an explicit element *inspireId*. Second, both elements remain independent this way – if deemed necessary a constraint could be established at some time requiring that the value of the *inspireId* be identical to the *coverageId*.

This establishes compatibility with CIS while, at the same time, retains semantics and use. There is one possible limitation, though, when using CIS 1.0: the type of *coverageId* is NCName which allows printable characters except colons (":"), as a GML heritage. CIS 1.1 liberates itself from this restriction and allows general strings.

In order to avoid introducing an extension over CIS 1.0 it was agreed at the corresponding 2017 INSPIRE Conference workshop to include the *inspireId* attribute within a metadata instance provided in the metadata bucket of the coverage. This is also useful for unambiguously referencing such metadata information to the coverage it describes.

Therefore, the best way to provide any additional information required by INSPIRE together with the standard coverage feature, is to include this information within the optional metadata bucket foreseen in the CIS 1.0 standard. With this aim, a metadata instance informed in this bucket shall adopt a theme-specific data type which includes all the additional INSPIRE elements required for each thematic domain.

In fact, this solution has been agreed in the INSPIRE context as a result of the online webinar 'Implementation of INSPIRE Coverages' held on 6th November 2017 [49], one of the last activities organized by the Thematic Cluster on Elevation, Orthoimagery, Reference systems and Geographical grids (see also [50–52]).

Factor out complex, but rarely used features

As can be seen aggregated or nested coverages complicate the user view on coverages: they need to deal with masses of different objects, each of which possibly contributes only some of its pixels. This is the traditional view of "1 coverage = 1 file" which eases implementers lives and complicates user lives. Seamless maps have proven handy since long, and with datacubes we find an extension of this principle to all spatio-temporal dimensions. That said, we acknowledge the necessity of such detail information for some expert cases, such as for ESA.

Therefore, we suggest to interpret the INSPIRE coverage specification as an INSPIRE Coverage Core and establish an INSPIRE Coverage Nesting Extension. This INSPIRE Coverage Core would have exactly one domain set, identical with its single, flat footprint therefore immediately mappable to the CIS domain set concept. Multiple extents and nested coverages are shifted into the more powerful, but also more complex extension. It is worth noting that such an approach is in perfect alignment with the modularization approaches adopted by both OGC and ISO. Figure 14 sketches this approach using standard UML notation.

This separation brings along a massive advantage for the theme designers, implementers, and users: when using only INSPIRE Coverage Core they do not need to think about the complexities of multiple, disparate extents and nesting, which significantly eases handling. Those users who actually want to make use of these advanced features will get exposed to it. For server implementers it allows to provide simpler implementations for the general majority of customers and provide an advanced package to those customers requesting it.

In passing we note that the concept of nested coverages, including modelling of the contributing footprints of each sub-coverage to the super-coverage as well as temporal validity, exist already in the WCS Earth Observation Application Profile (EO-WCS) [53], established earlier with funding by ESA. Hence, we recommend to align the INSPIRE modelling details in this respect with OGC EO-WCS, which offers even more powerful concepts plus streamlined request functionality, e.g., through dedicated request types.

Strictly map INSPIRE coverages to CIS

By now we are in a situation that allows mapping of the INSPIRE conceptual coverage model to OGC coverages.

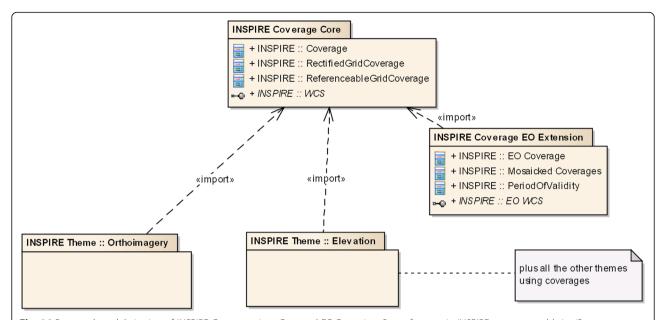


Fig. 14 Proposed modularization of INSPIRE Coverages into Core and EO Extension. Some features in INSPIRE coverages add significant complexity, but are rarely used in practice. Currently, an implementation (both clients and servers) need to implement the full plot. It is recommended to factor these advanced, more difficult features into one or more optional extensions which can be ignored where not needed, thereby easing the life of both implementers, data providers, and users in the common, simpler cases

Table 1 shows a possible starting point for such a mapping where the INSPIRE Core and Nesting Extension form the basis from which each INSPIRE theme can derive its own specializations, exemplified for Elevation and Orthoimagery coverages. Of course, such derivation needs to follow CIS rules for coherence and interoperability.

Bottom line, the INSPIRE Coverage Core can be mapped to CIS, actually to all versions: In CIS 1.0, the corresponding type is *RectifiedGridCoverage*, in CIS 1.1 type *GeneralGridCoverage* would be used with all-regular axes. We do not investigate further how this INSPIRE extension can be harmonized with CIS 1.1 because the multiple extents as well as nesting in INSPIRE leave open some design questions where the answers would be required for a rigorous mapping.

Re-harmonize INSPIRE with OGC

The path for achieving this objective implies the progressive adoption of the new OGC CIS v1.1, which provides a standardized way for encoding concepts foreseen in INSPIRE, like coverage aggregation (namely "coverage by partition" in CIS v1.0). Adoption of the new standard would boost the following benefits:

 Use of the most up-to date OGC standard for IN-SPIRE coverages, which does still allows previous versions of the coverage standards (GML 3.2.1, GML 3.3, CIS 1.0) while offering unifying concepts.

- Provision of a coverage architecture which is completely agnostic of any specific domain, and hence allows cross-domain use and combination of data
- Integration of regular and irregular grids from Earth observation, SOS, or other different thematic domains, with the generic concept of the CIS 1.1 GeneralGrid.
- Better integration of spatial and non-spatial domain sets, e.g. time series, or spatio-temporal coverages.

However, before adopting CIS v1.1 in the INSPIRE context there is a need to boost it in the software market, by promoting a wider implementation of the standard among WCS server and client tools by the respective vendors. Disposal of on-purpose funding for pushing actions in this direction would be extremely beneficial for boosting the data cube technologies and analysis potential, as well as to satisfy the urgent needs of relevant stakeholders in this domain area.

Additionally, a more complete analysis regarding IN-SPIRE coverages currently in place should be undertaken, taking into account the specificities from the different INSPIRE themes altogether. In particular, many of the properties considered in INSPIRE coverages, which are extensions to CIS and have to be distributed as coverage metadata, potentially overlaps with elements in the SWE Common Data model [15]. An exhaustive mapping is therefore necessary.

Conclusion

While raster data have long been neglected in defining Spatial Data Infrastructures they experience a massive increase in attention recently. This is partly due to the ever-growing deluge of data which become available more and more inexpensively, but also due to the great value of such data for society and business which is hard to overestimate. The appropriate concept for modelling 1D timeseries, 2D raster images, 3D x/y/t image timeseries and x/y/z voxel data, 4D x/y/z/t weather and climate data, among others, is given by coverages as standardized in OGC and ISO. Both are harmonized as ISO has adopted the coverage data model (CIS) and plans to adopt the coverage service model (WCS and WCPS).

The OGC/ISO coverage suite is currently enjoying a wide, increasing uptake. CIS with WCS is implemented by open-source projects like rasdaman, GDAL, MapServer, GeoServer, OpenLayers, and many more as well as by commercial products such as ESRI ArcGIS. User feedback is outstandingly positive, such as "Web developers who have not heard of OGC standards before immediately feel at home with these coverage standards" (Stephan Siemen, ECMWF) [13].

INSPIRE, acknowledging the foundational work accomplished by OGC and ISO, taps into this resource and reuses the coverage concept – however, with several modifications in the data model which may seem acceptable from a pure GML perspective, but effectively (i) lead to incompatibility with these standards and implementations based on them and (ii) add a unnecessary amount of complexity for all INSPIRE users while benefitting only a few experts. Consequently, generic WCS implementations can be expected to either break on INSPIRE coverages or silently ignore items.

Actually, OGC CIS and WCS standards provide clearly defined extension mechanisms which give flexibility while maintaining interoperability; these standards themselves contain already more than a dozen of extension sub-standards utilizing these mechanisms so their feasibility is proven manifold.

In this contribution we have made a comparison of OGC / ISO versus INSPIRE coverages, thereby spotting several incompatible divergences. We have also investigated the root causes, thereby finding a small set of patterns which are detrimental to interoperability. Based on this understanding we have been able to provide a small set of change proposals which (i) establish compatibility with OGC/ISO coverages, (ii) minimize changes to the INSPIRE coverage model, and (iii) simplify handling of INSPIRE coverages significantly.

We therefore recommend adjusting the INSPIRE definitions accordingly and perform extensive; the open-source rasdaman datacube engine [54], being official OGC Reference Implementation and widely used for scalable

spatio-temporal coverage services, seems a particularly suitable vehicle for this. This has been confirmed at the 2018 INSPIRE Conference where rasdaman was used for a quick prototyping of several coverage alignment proposals.

Additionally, we recommend a refactoring of the IN-SPIRE coverages definition in a simple, easy to handle mandatory Core covering most practical cases and optional Extensions for nested coverages (coverage aggregations), as well as for managing orthoimagery mosaics, etc. This extension should be harmonized with the OGC EO-WCS Application Profile standard which already addresses these concepts, and in a more comprehensive manner than pursued by INSPIRE. This is expected to massively simplify coverage handling for the large majority of day-to-day use cases.

Further, our investigation has been driven by the Elevation and Orthoimagery themes. Past experience show widely varying mileage concerning the theme experts' experience with modelling, using, and exploiting coverage data. Communities such as meteorology, with massive data storage requirements and data analytics needs, have explored datacube technologies over the last years; others, like mapping authorities producing Earth observation products, are just getting introduced in the subject, in terms of both deployment and how to leverage the benefits. Future work, therefore, should address all themes comprehensively so as to obtain a common, coherent, and dependable SDI framework. For a wider adoption across domains a wide range of data producers and consumers should be involved.

Concrete actions to be pushed and endorsed by the INSPIRE-MIG to achieve this cross-domain refactoring of INSPIRE coverages are:

- Revise the Technical Guidelines for the different INSPIRE themes, assuring consistency to CIS and presenting coverage data and services interoperability as an easily achievable goal.
- Urgently update the INSPIRE schemas for coverage data, which is at the moment the most important factor preventing interoperability of raster data in INSPIRE.

An additional avenue to be explored is how to embed coverages into different service specifications. While WCS offers the richest environment for coverage access, subsetting, and analytics there are more service APIs considered by INSPIRE, including WFS, SOS, and Atom. Currently, there is little exchange between the actors in these fields, and there is a need to strengthen cooperation with SOS and other domains on the one side and the coverage field on the other side, best by running specific practical activities the context of the INSPIRE Thematic Clusters.

The authors, together with an SOS expert, are currently preparing concepts for harmonization of coverages and Observations and Measurements (O&M) data so that such structures can be served through Web Coverage Service and Sensor Observation Service (SOS) simultaneously.

The goal and anticipated benefit of all these activities is making implementation of INSPIRE understandable and easier for data providers in the European member states. Additionally, it improving cross-theme and cross-cluster harmonization of INSPIRE coverages as well as increasing consistency with the encoding standards obviously is indispensable for data and service interoperability. Finally, a core goal is that existing, stable WCS implementations can be used. For vendors, this means they do not have to devise (expensive) special solutions, for users this means better interoperability.

Abbreviations

API: Application Programming Interface; ASCII: American Standard Code for Information Interchange; BLOB: Binary Large Object; CIS: Coverage Implementation Schema; CRS: Coordinate Reference System; DSM: Digital Surface Model; DTM: Digital Terrain Model; ECMWF: European Centre for Medium-Range Weather Forecast; EPSG: European Petroleum Survey Group; GML: Geographic Markup Language; GMLCOV: GML 3.2.1 Application Schema – Coverages (identical to CIS); ISO: International Standardization Organization; JSON: JavaScript Object Notation; MIG: INSPIRE Maintenance and Implementation Group; NCI: National Computational Infrastructure; OGC: Open Geospatial Consortium; RDA: Research Data Alliance; RDF: Resource Data Framework; SDI: Spatial Data Infrastructure; SOS: Sensor Observation Service; SQL: Structured English Query Language; WCPS: Web Coverage Processing Service; WCS: Web Coverage Service; WFS: Web Feature Service; WMS: Web Map Service; WPS: Web Processing Service; XML: Extensible Markup Language

Funding

This research is partially funded by the INSPIRE Thematic Clusters initiative, the European Commission under H2020 project LANDSUPPORT, and the German Ministry of Economics under project Big Data Cube.

Availability of data and materials

All data sets used in the discussion above are available from open, free, public sources as generally used in INSPIRE documents and discussions. Links are available from the authors on request. Data sharing not applicable to this article as no datasets were generated or analysed during the current study. XML code shown in this paper can be obtained from the authors by email.

Authors' contributions

PB contributed his OGC/ISO expertise in his capacity as editor of the. OGC/ISO datacube standards, JE contributed his knowledge about INSPIRE in his capacity of Thematic Working Group Facilitator EL and Ol. Together, they elaborated the assessment and remedial proposal. Both authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Large-Scale Scientific Information Systems, Jacobs University, Bremen, Germany. ²Head of Unit – Spatial Data Infrastructure of Catalonia (IDEC), Institut Cartogràfic i Geològic de Catalunya (ICGC), Barcelona, Spain. Received: 29 November 2018 Accepted: 14 January 2019 Published online: 11 February 2019

References

- United Nations. Transforming Our World. The 2030 Agenda for Sustainable Development. RES/70/1. http://www.un.org/en/development/desa/ population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E. pdf. seen 2018.
- United Nations. Sendai Framework for Disaster Risk Reduction 2015–2030. https://www.unisdr.org/files/43291_sendaiframeworkfordrren.pdf, seen 2018.
- United Nations. Summary of the Paris Agreement. http://bigpicture.unfccc. int/#content-the-paris-agreemen, seen 2018.
- N.n. Group on Earth Observations. http://www.earthobservations.org/index2. php, seen 2018-08-11.
- European Commission. COMMISSION REGULATION (EU) No 1089/2010, of 23 November 2010, implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial data sets and services (consolidated text). https://eur-lex.europa.eu/legal-content/EN/ TXT/PDF/?uri=CELEX.02010R1089-20141231&qid=1530548143791&from=EN, seen 2018
- EarthServer. EarthServer Big Datacubes at Your Fingertips. www. earthserver.eu, seen 2018.
- Baumann P, Hirschorn E, Maso J. OGC Coverage Implementation Schema, version 1.1. OGC document 09-146r6, www.opengeospatial.org/standards/ wcs. seen 2018.
- Baumann P. OGC Coverage Implementation Schema (CIS), version 1.0.1 formerly named GML 3.2.1 Application Schema - Coverages (GMLCOV). OGC document 09-146r2, www.opengeospatial.org/standards/wcs, seen 2018.
- Baumann P. OGC Web Coverage Service (WCS) Interface Standard Core, version 2.0. OGC document 09-110r4, www.opengeospatial.org/standards/wcs, seen 2018
- Baumann P. OGC Web Coverage Processing Service (WCPS), version 1.0. OGC document 08-068r2, http://www.opengeospatial.org/standards/wcs, seen 2018.
- 11. ISO. Geographic information Schema for coverage geometry and functions. In: ISO IS 19123; 2005.
- 12. ISO. Geographic information Schema for coverage geometry and functions Part 2: Coverage implementation Schema. In: ISO IS 19123–2; 2018.
- Baumann P. CoveragesDWG: OGC Spatio-Temporal Coverage / Datacube Standards. http://www.myogc.org/go/coveragesDWG, seen 2017.
- 14. Kottman C. Topic 6 Schema for coverage geometry and functions. version 7.0. In: OGC document 07–011; 2007.
- OGC. SWE Common Data Model Encoding Standard. http://www. opengeospatial.org/standards/swe, seen 2018.
- Baumann P. CRS Definition Resolver Web. http://external.opengeospatial. org/twiki_public/CRSdefinitionResolver, seen 2018.
- Whiteside A. GML Change Request for ReferenceableGrid. In: OGC document 07–112.
- OGC. OGC WCS 2.0 compliance test suite. http://cite.opengeospatial.org/ teamengine/about/wcs/2.0.1/site/index.html, seen 2018.
- Baumann P. The OGC web Coverage processing service (WCPS) standard. GeoInformatica. 2010;14(4):447–79.
- Liakos P, Koltsida P, Baumann P, Ioannidis Y, Delis A. A distributed infrastructure for earth-science big Data retrieval. Intl Journal of Cooperative Information Systems. 2015;24(2).
- Merticariu G, Misev D, Baumann P. Measuring storage access performance in Array Databases. New Delhi, India: Proc. 7th workshop on big Data benchmarking (WBDB); 2015.
- RDA. Array Databases: Concepts, Standards, Implementations. https://rd-alliance.org/system/files/Array-Databases_final-report_2018-03-19_0.pdf, seen 2017.
- 23. ISO. SQL 9075 Part 15: Multi-dimensional arrays (MDA). In: ISO IS 9075–15; 2018.
- 24. Baumann P, Rossi AP, Bell B, Clements O, Evans B, Hoenig H, Hogan P, Kakaletris G, Koltsida P, Mantovani S, Marco Figuera R, Merticariu V, Misev D, Pham BH, Siemen S, Wagemann J (2017) Fostering Cross-Disciplinary Earth Science Through Datacube Analytics. In. Mathieu PP, Aubrecht C: Earth Observation Open Science and Innovation Changing the World One Pixel at a Time, International Space Science Institute (ISSI), 2017.
- Karmas A, Tzotsos A, Karantzalos K. Scalable geospatial web services through efficient, online and near real-time processing of earth observation Data. In: Proc IEEE Intl Conf on Big Data Computing Service and Applications; 2015.

- Karmas KK. Scalable Geospatial Services for the Production of Time Series and Value-added Maps in Agriculture and Water Quality Monitoring. In: Proc. 2016 Conference on big Data from space (ESA BiDS'16); 2016.
- Marco Figuera R, Pham BH, Minin M, Flahaut J, Halder A, Rossi AP. Analyzing CRISM hyperspectral imagery using PlanetServer. Geophys Res Abstr. 2017:14632.
- Baumann P et al. Datacube services on a satellite: the ORBiDANSe project.
 Proc. international Astronautical congress, Bremen, Germany, 2018 (accepted).
- INSPIRE Drafting Team Data Specifications. D2.7 Guidelines for the encoding of spatial data v3.3. https://inspire.ec.europa.eu/documents/guidelinesencoding-spatial-data, published 2014.
- INSPIRE Drafting Team Data Specifications. D2.5 INSPIRE Generic Conceptual Model v3.4. https://inspire.ec.europa.eu/documents/inspire-generic-conceptual-model, published 2014.
- INSPIRE Drafting Team Data Specifications. D2.10.2 INSPIRE Data Specifications

 Base Models Coverage Types v1.0. https://inspire.ec.europa.eu/documents/inspire-data-specifications-%E2%80%93-base-models-%E2%80%93-coverage-types. published 2013.
- INSPIRE Thematic Working Group Elevation. D2.8.II.1 Data Specification on Elevation – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/ document/tg/el, published 2013.
- INSPIRE Thematic Working Group Land Cover. D2.8.II.2 Data Specification on Land Cover – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/ document/tg/lc, published 2013.
- INSPIRE Thematic Working Group Orthoimagery. D2.8.II.3 Data Specification on Orthoimagery – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/ document/tg/oi, published 2013.
- INSPIRE Thematic Working Group Geology. D2.8.II.4 Data Specification on Geology – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/ document/tg/ge, published 2013.
- INSPIRE Thematic Working Group Statistical Units D2.8.III.1 Data Specification on Statistical Units – Technical Guidelines v3.0. https://inspire.ec.europa.eu/ id/document/tg/su, published 2013.
- INSPIRE Thematic Working Group Soil. D2.8.III.3 Data Specification on Soil Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/document/tg/so, published 2013.
- INSPIRE Thematic Working Group Land Use. D2.8.III.4 Data Specification on Land Use – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/ document/tg/lu, published 2013.
- INSPIRE Thematic Working Group Environmental Monitoring Facilities. D2.8.III.7
 Data Specification on Environmental Monitoring Facilities Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/document/tg/ef, published 2013.
- INSPIRE Thematic Working Group Natural Risk Zones. D2.8.III.12 Data Specification on Natural Risk Zones – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/document/tg/nz, published 2013.
- INSPIRE Thematic Working Group Atmospheric Conditions and Meteorological Geographical Features. D2.8.III.13–14 Data Specification on Atmospheric Conditions and Meteorological Geographical Features – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/document/tg/ac, published 2013.
- INSPIRE Thematic Working Group Oceanographic geographical features and Sea regions. D2.8.III.15 Data Specification on Oceanographic geographical features – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/ document/tg/of, published 2013.
- INSPIRE Thematic Working Group Habitats and Biotopes. D2.8.III.18 Data Specification on Habitats and Biotopes – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/document/tg/hb, published 2013.
- INSPIRE Thematic Working Group Species Distribution. D2.8.III.19 Data Specification on Species Distribution – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/document/tg/sd, published 2013.
- INSPIRE Thematic Working Group Energy Resources. D2.8.III.20 Data Specification on Energy Resources – Technical Guidelines v3.0. https://inspire.ec.europa.eu/id/document/tg/er, published 2013.
- INSPIRE Maintenance and Implementation Group (MIG) Temporary MIG subgroup for action MIWP-7a. D2.9 Guidelines for the use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE v3. 0. https://inspire.ec.europa.eu/id/document/tg/d2.9-o%26m-swe, published 2016.
- OGC. Testbed-14. http://www.opengeospatial.org/projects/initiatives/ testbed14, seen 2018.
- OGC. OGC Spatio-Temporal Coverage / Datacube Standards. http://myogc. org/go/coveragesDWG, seen 2018.

- INSPIRE Thematic Cluster on Elevation, Orthoimagery, Reference systems and Geographical grids. Webinar Implementation of INSPIRE Coverages. https://themes.jrc.ec.europa.eu/pages/view/159283/webinar-implementation-of-inspire-coverages, dated 2017.
- INSPIRE Thematic Cluster on Elevation, Orthoimagery, Reference systems and Geographical grids. Workshop about Transformation of Coverage-Based Data Themes and WCS. https://themes.jrc.ec.europa.eu/pages/view/45690/ workshop-about-transformation-of-coverage-based-data-themes-and-wcs-barcelona-29-30-september-2015, dated 2015-09-29 and 2015-09-30.
- INSPIRE Thematic Cluster on Elevation, Orthoimagery, Reference systems and Geographical grids. Follow-up webinar on Coverage Data and Services. https://themes.jrc.ec.europa.eu/pages/view/77136/results-from-the-ws-on-transformation-of-coverage-based-data-themes-and-wcs-and-related-follow-up-webinar, dated 2016.
- INSPIRE Thematic Cluster on Elevation, Orthoimagery, Reference systems and Geographical grids. Workshop Implementation and potential of INSPIRE coverage data and WCS. https://themes.jrc.ec.europa.eu/pages/view/ 115418/results-from-the-ws-on-implementation-and-potential-of-inspirecoverage-data-and-wcs, dated 2016.
- Baumann P, Meissl S, Yu J. OGC® WCS Interface Standard Earth Observation Application Profile, version 1.0. OGC document 10-140r1.
- 54. N.n. rasdaman. http://www.rasdaman.org, seen 2018-08-11.

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com