ISO/OGC STANDARDS IMPACT ON NEXT GENERATION ENTERPRISE SPATIAL PLATFORM

Belay Teka Beshah[a], Max Martinez[a], Mladen Stojic[a]

ERDAS Inc., 5051 Peachtree Corners Circle, Norcross, GA 30092, USA[a]

Commission II, SS-15

KEY WORDS: Standards, Interoperability, Specifications, Decision Support, Information, Development, Integration

ABSTRACT:

The geospatial industry has finally caught up to the need for standardization required to manage the large amount of data and time consuming processes the community deals with constantly. The use of common standards is the most efficient way to achieve interoperability of geospatial information. Lately standards in the geospatial industry have come to maturity from the efforts of both the Open Geospatial Consortium (OGC) and the International Standards Organization (ISO). These standardization efforts enable enterprises to collect data once and to share it between different levels of the organization or across organizations. Moreover, the standards enable systems to be built that easily discover and seamlessly combine spatial information from different sources and share it among many users and applications. At ERDAS we have been building an Enterprise Spatial Platform that is designed from the ground up to securely manage and deliver geospatial data and on-demand value added information products to users operating on rich client desktop systems, web clients, and mobile clients. It was clear from the start that the success of an enterprise system like this depends upon the adoption of well established standards which led us to the investigation of all the relevant ISO/OGC standards. This paper summarizes the effort to understand, interpret, and implement the OGC and ISO standards pertinent to the geospatial community. It recaps the challenges faced in implementing the different ISO/OGC standards from a practical point of view.

1. INTRODUCTION

Achieving interoperability between disparate systems is increasingly becoming a focus point for organizations that deal with large amount of geospatial data. One approach to achieving these interoperability demands between systems is to employ evolving international standards, especially those circulated by the OGC and the ISO. Standards such as WMS, WCS, ISO 19115, CS-W, etc., provide the basis for the creation of geospatial systems that can interoperate with other system's components and does not lock the customer into the selection of a single vendor's solution.

The ISO and OGC specifications are mostly complimentary. Whereas ISO has mainly concentrated on abstract data modelling, the OGC has focused on interoperable services without specifying how the data is stored.

In the last couple of years we have also seen more customers, especially national, state and local government offices, driving the interoperability requirement by demanding that new systems out for bid support OGC/ISO standards for metadata and data delivery. The acceptance of these standards have protected the customer's investment by allowing them to switch from one system to another without a complete rewrite and have opened doors for vendors to compete on their implementation.

At ERDAS we recognized the business problem that people have in securely finding, describing, managing and delivering geospatial data to a multitude of client applications. We started building our next generation application on the premise that it is going to be a business system that knows about other systems in the organization and integrates well to solve business needs. The system consists of data crawlers for finding geospatial data and web services, harvesters for describing geospatial data according to metadata standards, and a geospatial data catalog with server side processing engines to create value added information products. It was clear from the start that the success of an enterprise system like this depends upon the adoption of well established standards which led us to investigate all the relevant ISO/OGC standards. We had to understand the rules for integrating the schemas from ISO and OGC standards into a consistent application schema. We discovered a lot of confusing points and duplication that made implementing a system that could adhere to the two sets of standards very challenging. We will reveal our scars from being on the forefront of the implementation of both published standards as well as some of the pending, committee draft, or final draft standards targeted for publication as standards in the near future.

In the following sections we will discuss the standards we dealt with extensively, devoting detail proportional to the effort required for analysis and impact on our next generation product.

1.1 ISO/OGC Implementations

Our involvement in the ISO and OGC standards went from the conceptual to the concrete. ISO/OGC are not concerned with specifying how the standards are to be implemented. However, in order to make implementations interoperable, a general structure and nomenclature must be specified.

We started by trying to conform to the Application Programming Interface (API) provided by the GeoAPI OGC working group. GeoAPI aims to reduce duplication and increase interoperability of Geographic Information Systems by providing neutral, interface-only APIs derived from OGC/ISO Standards.

One major advantage of the GeoAPI is that it acts like a bridge between software and the OGC / ISO standards process. The interfaces implied by the standards have been complimented where necessary by contributions from the open source community. Our acceptance of the GeoAPI interfaces was meant to allow our users to be able to mix and match different components based on standards from both commercial and open source implementations.

2. ISO GEOMETRY

At the core of any geospatial system is how to handle geometry. How to efficiently and correctly represent geometry has been a big discussion topic in the geospatial industry. Even within our organization we have multiple geometry representation modules that have grown organically through the years from different programs and acquisitions. Some of the issues we were trying to address were:

- Redundancy: In previous systems, we ended up with too many different ways of expressing the same geometric concept.
- Simplicity: We desired an API that provides simple utilities to keep developers from making common mistakes.
- Interoperability with other systems: We need to provide convenient support for interfacing with other APIs.
- Coordinate Reference System Support: Geometric objects cannot be properly related without an associated coordinate reference system.
- Spatial Operators: support for simple operators like Disjoint, Intersects, Contains, Unions, Relate, etc.
- Advanced Features: We require expandability for complex topological operators to do various kinds of vector processing.

2.1 Implementation

There are currently different widely used geometry models and implementations like Java 2D/3D, Java Topology Suite (JTS), GeoTools, etc. After evaluating these specifications we felt that the ISO spatial schema addressed most of our requirements. ISO 19107 is a very thorough model of spatial characteristics and few applications will require its full range of capabilities. As we started looking in earnest we quickly realized that we do not need all the functionality. We concentrated on a few essential classes and behaviors: Point, Line (Curve), and Polygon (Surface) with their spatial operations. We worked together with the open source community to upgrade the GeoAPI interfaces to take the full ISO geometry specification into account. Our collaboration with Refraction Research resulted in a Geometry package in GeoTools that implemented initially only a very thin line through the specification. This module went through sufficient conformance and coverage testing in order to satisfy our confidence in its correctness by reusing the JTS test harness.

3. ISO COVERAGE

Coverage associates positions within a bounded space (its domain) to feature attribute values (its range). Some common examples of coverage include a raster image, a digital elevation model, or a temperature observations. A simplified way to think of coverage is a model that provides data like raster and terrain about some area bounded in space and time.

The first type of dataset we wanted to model using the ISO coverage concept was a raster image. An image was modeled as a DiscreteGridPointCoverage from ISO 19123. This led to a lot of discussions from how to reuse our existing libraries to support the coverage model to the simpler question of what do we return as self describing range values for each domain location. The basic operation of getting raster and terrain data about an area bounded by space-time was a real challenge even though it was a matter of implementing a single “select” method on the coverage API that is applicable to all kinds of coverages. Moreover, the GeoAPI interface definition of ISO 19123 posed a logistical problem due to the fact that there was an existing coverage interface definition based on the OGC's grid coverage specification. At the time of our investigation, the OGC specification was planned to be retired in favor of the ISO 19123 based coverage specification. However, there was no progress on this front from the community except a couple of investigations (Nordgren, 2006) and some noise in the open source geospatial community mailing lists.

ISO 19123 is one of the most complicated specifications that we came across. In theory it provided the most extensive abstract modeling, allowing both continuous and discrete observations to be treated in a self describing manner. This is a key characteristic for establishing interoperability between systems. However, there were a lot of problems due to inconsistencies in its usage of other specifications. Efforts by other users to model ISO 19123 have revealed similar difficulties as well as performance concerns (Nordgren, 2006). In addition, ISO coverage is a very cumbersome interface for clients to work with for accessing imagery or gridded data as a block. Block oriented access is the de facto access pattern used by imaging applications. We finally came to the conclusion that, it is better if the ISO Coverage API were one of the ways to access gridded data but definitely not the major gateway to access gridded data using our libraries or server because of its inefficiencies.

4. ISO METADATA

Metadata has been a hot topic for many years not only in the geospatial industry but also in the IT industry as a whole. It has been considered one of the most complicated aspects of the information age since we have been overwhelmed with data. Finding the right data for the job at hand has become increasingly difficult. Metadata could be simply defined as "data about data.". Used in the context of digital spatial data, metadata is the background information which describes the content, quality, condition, and other appropriate characteristics of the data.

Expressing these characteristics of data in a standard way that both machines and humans can understand has been a challenge. Many organizations have tried to come up with a way to achieve this with some level of success. For example, the Federal Geographic Data Committee (FGDC) metadata model has seen extensive usage in the US government and commercial entities that are closely related to government, but has struggled to gain acceptance internationally.

Here is where the ISO Technical Committee ISO/TC 211, Geographic information/Geomatics has taken the lead and provided the metadata standards 19115, 19115.2 and 19119. These standards are detailed information models for describing datasets and services and have been very successful by concentrating on what needs to be defined and not how it is stored. This allows even legacy systems to be able to produce these information models from the disparate physical storage models. ISO 19139 on the other hand was specifically built as a physical model that allowed different systems to exchange metadata. By building on top of a very well defined XML schema it has allowed tools to be easily built that could extract the information from XML files or streams.

4.1 Implementation

When we started implementing the interfaces from GeoAPI for metadata we found a lot of confusion and inconsistencies in the API. This lead to our effort to synchronize the API's published in GeoAPI with the latest correction to the specification ISO 19115:2003/Cor.1:2006(E). At the same time, we attempted to incorporate modern Java language features, such as generics, in the API's requiring change to make them easier to use. In our partnership with Refraction Research to accomplish these tasks more than ten bugs and feature enhancements were submitted to the bug tracking system of GeoAPI. These submissions resulted in more than fifty changes in interfaces or interface signatures. By going through the rigorous change request process we were able to solicit a lot of feedback from the community. We also took the extra step required to make sure our changes are beneficial to the community at large by investing time and resources in modifying GeoTools metadata packages to conform to these API's in addition to ours since it is one of the reference implementations.

By relying on the standard we were able to build one metadata harvester that is based on the ISO 19139 specification and uses XSLT(which is a language for transforming XML documents into other XML documents) to translate many other metadata standards to the ISO 19139 conforming XML. This resulted in an automatically validated metadata import into our catalog. The acceptance of the metadata standards 19115 and 19139 has resulted in many vendors providing the option to get the metadata for their data products in this format or providing tools or translation to ISO.

5. ISO IMAGERY SENSOR MODELS

The Committee Draft(CD) ISO 19130 Geographic Information – Imagery Sensor Models for Geopositioning is arguably a sine qua non for an enterprise spatial platform that wishes to deal interoperably with imagery that is not yet geo-rectified. This specification attempts to define an abstract data model for imagery sensor models. In its current form, this includes four basic physical sensor types: frame, pushbroom, whiskbroom, and synthetic aperture radar (SAR). As ambitious as this is, the specification also attempts to model other common means of relating a coordinate reference system which is supported by an image or engineering datum (as would be initially associated with a remotely sensed dataset) to a coordinate reference system supported by an earth-based datum. This includes representations of true replacement models, correspondence models, and ground control points.

Understanding the necessity of such a broad model is aided by considering a layered model for describing (quadrilateral grid) coverages (Baumann, 2003). Coverages, of which an image dataset is one kind of example, are the concern of ISO 19123 and the OGC's Web Coverage Service implementation specification.

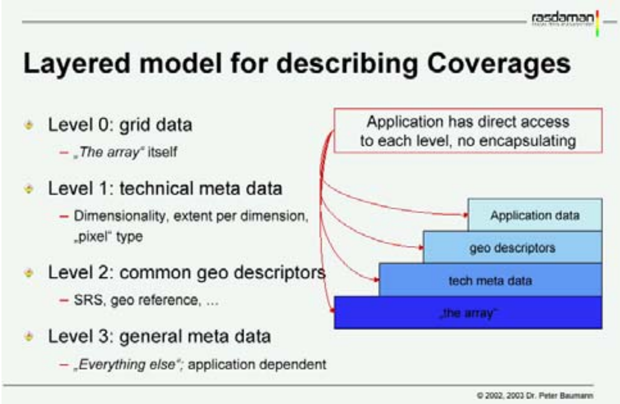

Figure 1. Layers of Coverage Data

ISO 19130 addresses the Level 2 information in a quadrilateral grid coverage that has not been geo-rectified. The descriptive name for this Level 2 information might better be phrased as "essential geo descriptors", because without this information, the dataset cannot be related to other geospatial data, and thus is of limited use to an enterprise spatial platform that seeks to deliver on interoperability.

That said, as we progress through increasingly complex forms of geo descriptors, there is a corresponding shrinking of concerned audience. Specifically, the audience for georectified imagery is much larger than the audience for georeferenceable imagery which is larger still than the audience seeking to refine image geolocation models.

In the immediate future, the creation and refinement of geolocation models is likely to continue to require the expertise and understanding of a small audience. Georeferenceable imagery, on the other hand, suffers from two impediments to wider spread use, both of which seem to be within the grasp of manageability: the interoperability of coordinate transformation descriptions, and the speed of performing on-the-fly image resampling through these complex coordinate transformations.

ISO 19130 promises to solve the former interoperability problem.

5.1 The Form and the Content

To be successful in attempting to standardize an abstract data model for geolocation in the context of a larger set of standards, ISO/CD 19130 must present a form that can be conceptually harmonized with other standards and deliver content that can allow geolocation to be performed properly by different implementations.

5.2 The Form of Geolocation Information

The form of geolocation information includes not only how it is presented as a concept but also (and almost more importantly) where and in what form that concept is relatable to other standards.

ISO 19130 is currently attempting to hook into two other standards, ISO 19115 and ISO 19111, through the concept of geolocation information.

The geolocation information concept of ISO/CD 19130 is related to ISO 19115 via another draft standard ISO/CD 19115-2 Geographic Information – Metadata – Part 2: Extensions for imagery and gridded data. In the current approach, the MD_Metadata of ISO 19115 discoverable in a dataset has an association to an MD_SpatialRepresentation, an abstraction which has an MD_GridSpatialRepresentation specialization for gridded datasets. The MD_GridSpatialRepresentation is further specialized by being either MD_Georectified or MD_Georeferenceable. At this point ISO/CD 19115-2 takes over and extends the MD_Georeferenceable concept with its own MI_Georeferenceable concept so that this georeferenceable spatial representation is associated to zero or more sets of MI_GeolocationInformation. This MI_GeolocationInformation is then specialized within the ISO/CD 19130 specification. This is a reasonable approach and will sit comfortably with the small audience of experts that create and refine geolocation models.

The hook that has the potential to reach a much larger audience, however, is the hook into ISO 19111. Yet there are a couple of problems standing in its way.

The first problem is that the relationship is not developed enough. This relationship has only begun to be developed in the most recent draft of ISO/CD 19130 that we have seen (10-19-2007), and it only consists of additional classes and relationships in a UML diagram without any accompanying narrative. In this draft a new class, SD_Transformation, is subclassed from the CC_Transformation of ISO 19111 as well as from something ISO/CD 19130 calls an SD_CoordinateGenerator. The SD_CoordinateGenerator is itself a specialization of MI_GeolocationInformation. Along the way these additional classes and their subclasses are associated to all the other specializations of MI_GeolocationInformation.

This brings us to the second problem, the CC_Transformation that organizes the MI_GeolocationInformation should be most naturally accessed not through the metadata model of ISO 19115, but through the coverage model of ISO 19123. Any coverage (not just quadrilateral grid coverages) whose coordinate reference system is supported by an image datum or engineering datum requires the essential geo descriptor of a georeferencing transformation to be used in a geospatial context with other datasets. Although OGC WCS seems to have this mostly right through the current service interface, ISO 19123 does not, itself, incorporate the CC_Transformation of ISO 19111 in its model. ISO 19123 banishes (what would be) the ISO 19111 CC_Transformation to the black box of the behavioural interface of the CV_ReferenceableGrid. This is no place for information that is striving to be interoperable. Interested OGC participants are currently examining this issue as it naturally impacts the discovery of this information in the model as well as the proper encoding of the model in Geography Markup Language (GML).

Use of MI_GeolocationInformation in a CC_Transformation has broad applicability. Thus both the work of defining this relationship as well as the work of positioning a CC_Transformation properly in ISO 19123 should be completed in order for ISO 19130 to achieve its potential as a standard.

5.3 The Content of Geolocation Information

The primary difficulty with the content of geolocation information is getting all the details just right and clearly explained. With something as complex as image sensor geometry models, this is unlikely to be achieved through specification alone. Implementation will need to be developed to determine what was incompletely specified. Additionally, it is often found that the specification is complete given idealized datasets, but datasets are rarely ideal in the real world.

For instance, in our experiments with the SD_TrueReplacementModel form of MI_GeolocationInformation, we find that although generated from a physical sensor model at one point in time, this association (to the physical sensor model from which it was generated) is often lost in real world datasets. It then becomes impossible to present the dataset's available information using the SD_TrueReplacementModel portion of the ISO/CD 19130 model because there is no way to convey what the intended ground coordinate reference system of the geolocation information is.

Additionally, the SD_CorrectionTable that initially served the Universal Real Time form of the True Replacement Model from early ISO/CD 19130 drafts needs to be properly updated to accommodate the TRM forms discussed in later drafts (image or ground side correction via polynomial function).

5.4 The Sensor Web Enablement (SWE) Dimension

Complicating these types of harmonization and specification completeness difficulties is the arrival of SWE on the OGC scene. The SWE initiative delivered a set of specifications that offer delivery and description of data that is much closer to the sensor than the data delivered by WCS and WFS. Thus SWE has its own manner of delivering geolocation information in the form of a SensorML ProcessModel.

There seems to be no current best practice for where this information is to be found in the larger model. The standards bodies should likely strive for making it accessible via the coverage result of an observation through the Sensor Observation Service (SOS). This will achieve consistency between the ISO and OGC standards if and when ISO/CD 19130 gets its relationship to ISO 19111 sorted out and WCS and ISO 19123 come to agreement on the correct model for delivering a georeferencing transformation with a coverage.

Harmonizing the content defined in ISO/CD 19130 with the content of SensorML ProcessModels describing georeferencing transformations, however, will likely pose more difficulties. SWE's ProcessModels are coming to life as practical working models. Initial models are focused on the (more widely applicable) georeferencing transformation itself and leave the topics of error propagation and refinement unaddressed. As these models gain traction in working implementations, it becomes increasingly important for the SWE and ISO/CD 19130 committees to be working on the harmonization of these models so that the practical approach of SWE serves to compliment the specifications approach of ISO/CD 19130. We do not want to end up with two competing standards and double the mental effort we all need to apply to this already mentally taxing problem.

6. OGC CATALOG SERVICE

The Catalog Service from OGC defines an open standard for modeling, publishing and searching metadata for data, services, and related information objects. To support its interoperability goal the catalog service specification defines a minimal common query language (the OGC Common Catalog Query Language) which allows disparate systems to share a common catalog.

6.1 Specifications

For many historical reasons, Catalog-oriented specification at the OGC is split into several documents that address different aspects. That is the main reason why the Catalog specification seems quite complex to people not-involved continuously in the catalog working groups.

6.1.1 OGC Catalog Services Specification

This is one of the more abstract specifications, explaining the behavior of a catalog, and defining the basics for CS-W (HTTP), CORBA and Z39.50 implementations. CS-W is the most discussed and probably the more used Protocol Binding. It seems that CORBA and Z39.50 are not currently active and are present in the specification for historical reasons.

6.1.2 OGC ISO Metadata Application Profile of CS-W

This Application Profile (AP) for CS-W uses the ISO 19115/19119 Information Model to store ISO 19139 and ISO19119 Metadata files. Unlike the ebRIM profile, this AP is not 'generic' and only allows ISO Metadata cataloguing.

6.1.3 OGC ebRIM Profile of CS-W

This Application Profile for CS-W uses the ebXML Registry Information Model (ebRIM) metamodel as 'generic' Information Model to store and retrieve any artifact in the Catalog. It is generic in the sense that ebRIM objects allow representing any complex structure (OGC Web Services, ISO 19139 Metadata files, CRS, Earth Observation Products ...) with the cost of defining a model (what we call an 'Extension Package') for any of these artifacts.

In 2007, the OGC strongly endorsed the ebRIM profile as the preferred basis for future profiles of OGC catalog, encouraging communities to define and develop a variety of ebRIM Extension Packages, standardized in the OGC, for various application domains. Some of these extension packages are:

- **Basic Extension Package**: defines a set of predefined objects that all compliant CS-W ebRIM Catalogue must support.
- **EO Products Extension Package**: describes how the Earth Observation products metadata is modeled and organized for Heterogeneous Earth Observation Missions Accessibility (HMA).
- **ISO Metadata (CIM) Extension Package**: used for the discovery and management of ISO 19115 and ISO 19119 metadata. This Extension Package was intended to allow CS-W ebRIM catalogs to provide the same functionalities that CS-W ISO Application Profile catalogs do.

6.2 Implementation

Obviously, the many choices available for catalog implementations cause interoperability issues (i.e., it is not possible for a CS-W ebRIM catalog designed to store CRS to really interact with a Z39.50 Catalog or even with a CS-W ISO 19115 AP one).

Through the years, there have been many debates in the working groups about ebRIM vs. ISO Application Profiles. In addition, all these minor inconsistencies and degrees of freedom in the different documents cause confusion. In its latest acceptance of ebRIM as the preferred cataloguing meta model, the OGC did not take that extra step required to unambiguously choose a catalog specification profile. Instead it still left the door open for the other specification to continue development.

The different specifications are not only none interoperable from a practical user point of view but also result in a lot of duplication of effort for catalog service implementers. Even within the ERDAS family of products, with our current acquisitions we have multiple implementations of the catalog. Having a common application specification that is designed with one and only one unambiguous API whereby vendors compete by quality of implementation would be highly beneficial. Hopefully, the OGC Catalog community is aware of these problems and all OGC Revision Working Group are working to address these issues.

7. OGC WEB MAP SERVICE

The Web Map Service defines a simple interface for web based mapping applications to produce and consume a map of spatially referenced data dynamically from geographic information using a standard image (GIF, PNG, JPEG, or other format). The OGC WMS specification is one of the most successful specifications in the geospatial community. At last count there were around 194 Compliant (OGC Implementation, 2008) commercial and open source implementations. Its ease of use for client applications that range from simple web site to complicated 3D viewer has resulted in a large number of server implementations of this specification. Its unequivocal nature of client-server interaction has greatly contributed to its success.

If there is one criticism that could be levied against the WMS it is that this service is not well suited for fast delivery of large amounts of data. While it is good at delivering a well formed map for various clients it simply is not responsive enough for highly interactive work. For single datasets, complementing WMS with emerging standards like JPEG 2000 Internet Protocol (JPIP) might solve the performance problems by providing direct access to JPEG 2000 image compression algorithm's scalability by allowing a client and server to negotiate for the delivery of only portions of the image file.

8. OGC WEB COVERAGE SERVICE

The Web Coverage Service (WCS) supports the exchange of raw geospatial data as "coverages" that are bound in space and time without the need to apply portrayal.

There are not as many clients that consume WCS as are available for WMS. However, rich remote sensing client like ERDAS IMAGINE need a WCS server to provide access to raw data so that a user will be able do analysis or interpretation of data on the fly. Most of the WCS web clients we found were used to allow users to select the area they are interested in and download data to be used in rich desktop applications. As time goes on we believe there will be a real intake of WCS as people move from precooked static maps to real data to be used in analysis and information extraction.

9. OGC WEB MAP CONTEXT

A Web Map Context (WMC) is an XML document that describes the appearance of layers from one or more OGC compliant services (WMS, WCS, and WFS) and can be transferred between clients. It is a portable platform-independent format of storage while maintaining startup views, the state of the view and storing additional layer information.

The Web Map Context has been very useful in allowing us to bridge the gap from our legacy desktop application to the services provided by the enterprise servers. Our ability to load and save WMC files from both our desktop products and web clients have facilitated the reuse of our components. More importantly it has created a standards conforming way of exchanging data between our products by allowing us to avoid creating another format of our own. This was considered when we started investigating how to allow our desktop products to connect to our server or actually any standard conforming geospatial service (WMS, WCS, CS-W).

10. CONCLUSION

Being at the forefront on implementing these standards has exposed us to a lot of pain points but has allowed us to take advantage of the thinking, planning, and design that others have already done.

It is clear to us that the standardization process can only proceed successfully when coupled with reference implementations that resolve problems that may not be apparent through specification alone.
It is also obvious that harmonization between specifications dealing with different matters is necessary once independently developed topics begin being adopted by a single large system or multiple systems

The advantages of standards adoption, however, are worth the trouble in the end. For example, following these standards our server side products have been used immediately in unanticipated ways. One of the most exciting things for us has been how our customers and integrators have been able to take existing clients and modify them to use our servers within days of our products being released.

REFERENCES

Baumann, P., 2003. Layered model for describing Coverages, Munich, Germany. https://portal.opengeospatial.org/twiki/bin/view/WCSrwg/PeterBaumannCoverageMetadataHierarchy (accessed 25 Apr. 2008)

Nordgren, B., 2006, An ISO19123 Coverage Primer,. http://docs.codehaus.org/display/GEOTOOLS/ISO+19123+progress+and+future+plan (accessed 25 Feb. 2008)

OGC Implementation, 2008, OGC Implementations by Specification, http://www.opengeospatial.org/resource/products/byspec/?specid=168 (accessed April 2, 2008)

Skelton, B., Stojic, M., Strynatka, R., 2007, Photogrammetric Week 2007, pp. 257-267