InfraGML Proposal (13-121)

OGC Land and Infrastructure DWG/SWG

Paul Scarponcini

November 5, 2013

**1. Introduction**

The Open Geospatial Consortium (OGC) Land And Infrastructure Domain Working Group (LandInfraDWG) was chartered to “focus on determining how best to integrate and support the LandXML schema within the OGC framework, as well as how to better manage and integrate CAD-based land information with other OGC standards” [1].

The Problem Statement in [1] goes on to say that

the LandXML data format is currently not being supported by a standards organization. This has resulted in confusion in the marketplace as to the future of the standard as well as the fact that there has not been any work done to advance the standard since 2009. LandXML is used worldwide as a neutral data exchange format by a number of government agencies and private sector firms to share land development, civil, survey and other infrastructure-related data. LandXML data is of value to the larger geospatial data community, but currently the format is not integrated with any of the OGC standards. Both domains would benefit from integrated access to this information.

The Charter commits the LandInfraDWG to the following activities:

1. Reviewing the current LandXML schema and determining how best to continue to support the existing users and engage with them.
2. Assessing the current industry support for the LandXML schema and whether multiple, incompatible versions of the schema have evolved.
3. Investigating how to best incorporate LandXML into the OGC standards framework, including identifying places where there may be common elements in existing standards such as CityGML.
4. Investigating the possibility of moving the LandXML schema into LandGML.
5. Other efforts to integrate land information contained in various CAD formats into the OGC standards framework.

The Charter makes the business case for this undertaking by acknowledging the LandXML user community as consisting of over 650 organizations with 750 members in over 40 countries and support by over 70 registered software products. Because of the loss of active leadership and development over the past five years by LandXML.org, this community is currently unsupported. “By OGC adopting the LandXML schema, the user community will be assured of a formal process for maintaining, improving, documenting and in fact, formalizing the standard. This will lead to greater confidence in use of the standard and new opportunities for integration with other related OGC standards.”

One of the initial problems for the LandInfraDWG to solve was to gain a better understanding of exactly what LandXML is and does. LandXML-1.2 contains almost 5000 lines of minimally documented XML code covering some 16 subject areas. There is no formally published documentation, user guide, requirements definition, or underlying conceptual model.

The Land and Infrastructure Standards Working Group (LandInfraSWG) was therefore chartered [2]. Its first task was to develop a UML as-is conceptual model of LandXML-1.2 to:

1. Aid in the understanding of what LandXML 1.2 is and does
2. Provide the basis for a conceptual model of what a future Land standard should do, based on an assessment of user requirements
3. Establish a single set of consistent concepts that can be implemented in any set of implementation-specific standards, such a LandXML 2.0, a LandGML, a LandSQL, etc.

The LandInfraSWG was chartered as a persistent SWG to enable it to work on:

1. a conceptual model for land development,
2. subsequent potential land information and development implementation-specific standards proposed by the Land And Infrastructure DWG.

**2. LandXML-1.2 Assessment**

Work began on reverse engineering a UML conceptual model for LandXML-1.2 in November, 2012. By June, the Units and Alignments subject areas were modeled and documented [3]. Several problems with the LandXML-1.2 were discovered and likewise documented. Work continued on the Coordinate System, Pipe Networks and Survey subject areas. Similar problems were encountered in these areas as well.

Problems discovered with LandXML-1.2 included the following [3]:

1. **Optional Content**: The apparent goal was flexibility as opposed to interoperability. The result is a comprehensive set of properties, but most of them are optional to allow individual users to choose which ones to implement. LandXML can therefore more appropriately be characterized as a negotiable template rather than an interoperable standard.
2. **Schema Document Structure**: The schema document is not organized in a structured manner, such as by logical packages. It is therefore difficult to find elements without electronic search, for example to determine which types are used within multiple element definitions. Modifying a type could have inconsequential impact on other parts of the schema.
3. **Choice Compositor:** The schema has some serious problems with choice compositors. The intent of a choice compositor is to allow an XML document to contain a single element from the choice list. By making maxOccurs unbounded for the choice compositor, any number of choices can be made. This does not mean that any number of elements on the list may be selected for a single choice. For each (mutually exclusive) choice, a single element from the choice list is selectable. In many cases, this is not the intended outcome, for example, a Spiral must have three choices (minOccurs="3" and maxOccurs="3") between Start, PI, and End but because a choice compositor is used, three Starts and no PI’s or End’s would be a valid solution. If any element on the choice list has a cardinality of [0..1], the result is that this element can be chosen, but because it is optional, it is possible to supply no value – even if the choice itself has a minOccurs of one.
4. **Station Data Type**: There is a Station type derived from Double. However, there are numerous situations where an element uses Double as its type instead of Station. For type Station, it is clear that the value specified is the internal station, that is, the absolute distance from the start station value having already factored in station equations. Double could be the unadjusted station value occurring after a station equation and is therefore ambiguous.
5. **Weak Point Typing**: Point types are too weakly typed. A Point is defined as “a text value that is a space delimited list of doubles.” It is usually used in a traditional sense as point coordinates in the form of “northing easting”. However, it can also be used in much less obvious cases where the meaning of the coordinates depends on the context of how the “point” is used. In the Profile package, the ProfSurf PntList2D point coordinates represent station and elevation. In the CrossSections package, the CrossSectSurf PntList2D point coordinates represent either offset distance and absolute elevation or per cent slope and horizontal distance. In the latter case, the point is not a point at all. It is a line segment.
6. **Name Optionality Inconsistency**: Though most elements have name and desc or description attributes, there does not appear to be consistency in their optionality. Typically, for a collection (e.g., Alignments) name is optional and for those things in the collection (e.g., Alignment) name is mandatory. This is not always the case – name is optional for both CrossSects and CrossSect.
7. **Case Inconsistencies**: Conventions for the use of Upper- vs. lowerCamelCase are generally but not consistently adhered to as with OGC and TC211 standards.
8. **Confusing Plural Names**: When an element is a set of another element type, the name of the set is just the plural of the member element name. For example, Alignments is a set whose members are of type Alignment. But this is not always the case. Profile (in the singular) is actually a set of profiles.
9. **“Unique” Identifiers**: The inclusion of an object ID (oID attribute) does not appear to be consistent or explained. For many elements, a name is allowed instead, though not required and not guaranteed to be unique. They were apparently added as an afterthought, in support of user applications that include such unique identifiers.
10. **Feature**: The LandXML notion of Feature as an attribute is upside down from a feature-based model as specified in ISO TC211 and OGC spatial standards. In LandXML, Features are attached to element classes instead of attaching elements to Features. In LandXML, Feature is a user-definable property (and its value) which can supplement the element definition. But many times this is one of the choice compositor’s choices, disrupting the count of meaningful selections. In some cases (e.g., Monuments), a collection is specified as containing one or more subject elements followed by zero or more features. It is not clear if (and how) the individual subject elements are meant to align with the individual features.

Fixing many of these problems will break backwards compatibility. For example, changing the choice compositor to the sequence compositor it should have been will introduce forced ordering of the choices which may not have been followed by previously valid XML documents. Increasing typing strength and increasing mandatory content would have a similar affect.

Because LandXML has not been updated for over five years, it has failed to keep up with newer civil design methods and software. Three dimensional modeling, for example, is now more prevalent. Keeping pace with these changes will necessitate additions or modifications to the LandXML schema.

The following issues highlight LandXML’s disparity with the OGC standards baseline:

1. **Modular Specification**: LandXML-1.2 does not follow the OGC Modular Specification. There are no discernible modules, no formal requirements defined, no mandatory core functionality specified, and almost everything is optional.
2. **Conceptual Model**: There is no conceptual model defining key LandXML concepts, their definitions, properties, or associations. It is inconsistent with the concepts specified in the OGC Abstract Specifications. It is therefore difficult to understand exactly what is called for by the XML schema.
3. **Feature-based**: LandXML does not support the feature-based model of OGC and TC211 standards. Though it does support “features”, its notion of what a feature is is completely different than that of OGC.
4. **Geometry Model**: LandXML does not use the harmonized geometry model supported by OGC, TC211, and SQL/MM standards and specified in ISO 19107 (OGC Abstract Specification Topic 1). It may be possible to map some of the geometry types between LandXML and the harmonized model.
5. **GML**: LandXML uses W3C XML rather than OGC geospatially extended XML (GML).
6. **Case Conventions**: LandXML does not consistently support the use of lower- and UpperCamelCase for naming as established by OGC and TC211.

**3. InfraGML proposal**

To be able to support LandXML as an interoperability standard, it would be necessary to fix the problems identified above. As this would break backwards compatibility, it would need to be version 2.0. Unfortunately, too much of the thinking behind LandXML-1.2 has been lost due to the lack of documentation and conceptual modeling. A LandXML-2.0 would also need more up-to-date functionality. We are unable to determine a list of contributors, users or supporting software applications in order to determine which parts of LandXML are still relevant.

Therefore, a fresh start standard is proposed having a use case driven subset of LandXML functionality, but implemented with GML and supported by a UML conceptual model. Called InfraGML, this new standard would:

* be supported by a recognized, dependable Standards Developing Organization, OGC
* align with existing OGC (and TC211 and SQL/MM) standards, including the modular spec
* benefit from functionality already supported by GML, including features, geometry, coordinate reference systems, linear referencing, and surface modeling (TIN)
* initially focus on survey and alignments, the two subject areas which have identified need and committed resources for development
* using modular extensions, be able to expand into other areas (e.g., pipe networks, parcels, etc.) as resources become available
* be use-case driven
* be based on a UML conceptual model developed prior to GML (and any other future) encoding
* have more up-to-date functionality
* be synchronized with the concurrent efforts by buildingSMART in their development of Infrastructure-based Industry Foundation Classes (IFCs)
* be more easily integrated with CityGML and TransXML

The alignment part of InfraGML would overlap the recently announced IfcAlignment development work of buildingSMART International (bSI). The full facility life cycle list of use cases to be supported by InfraGML would be broader than the IfcAlignment focus on design and construction. OGC and buildingSMART can work together on the use case definitions and conceptual model so that the two resultant implementations (IFC and GML) would be harmonized. The resultant IFC version would be consistent with the buildingSMART standards baseline and the GML version would be consistent with the OGC standards baseline. This would enable interoperability within each respective baseline (e.g., IfcAlignment with COBie and InfraGML with CityGML). The shared conceptual model would allow cross-baseline interoperability. When presented to bSI at their 2013 Munich Summit, this proposal was well accepted by their Infrastructure Room (formerly OpenINFRA). At the Summit, they passed the following resolutions [4]:

|  |  |
| --- | --- |
| IR 13/10:7 | The Infra Room resolves to give top priority to the Alignment project as expressed in the project summary of IFC Alignment and decides that alignment is needed in the next version of IFC. The first task of Infracom is to prepare a project management plan. |
| IR 13/10:8 | The Infra Room concludes that coordination and collaboration with OGC and LandXML is a crucial issue for the alignment project. |

**4. Proposed Actions**

If accepted by the Land and Infrastructure DWG and SWG, a new project within the SWG should commence immediately to develop the proposed InfraGML standard in conjunction with the buildingSMART International Alignment Project. The following steps are proposed:

1. **Subject area identification**: The list of initial subject areas for InfraGML shall be identified. Additionally, other areas may be identified as possible future extensions.
2. **Use case definition**: Use cases relevant to each subject area should be identified and described. This should be harmonized with the bSI Alignment Project use cases, the building SMART Alliance (bSA) BIM-GIS use case list, and any forthcoming use case definitions from the ISO TC211 GIS-BIM ad hoc committee.
3. **Conceptual modeling**: A UML conceptual model of the initial subject areas shall be developed, including a separating out of Core functionality. This model shall be harmonized with the model developed by the bSI Alignment Project.
4. **RFC**: A public Request for Comment shall be issued based upon the UML conceptual model, prior to any GML encoding. This is to solicit input from the existing LandXML community (and others), including users as well as software developers. The conceptual model shall then be revised accordingly.
5. **GML encoding**: A GML 3.2.1/3.3 compatible standard shall then be developed based on the revised UML conceptual model.
6. **Extensions**: Similar steps shall be followed for subsequent subject areas.

 As with all voluntary OGC efforts, the timing of these steps will be dependent on the resources which step forward to contribute. Every effort shall be made to focus on those subject areas where resources are available in order to keep in step with bSI.

**5. References**

[1] Burggraf, David, Scott Simmons, Hans-Christoph Gruler and Gene Roe, *Land and Infrastructure Domain Working Group Charter*, OGC Project Document 12-079r3, August 20, 2012, <https://portal.opengeospatial.org/files/?artifact_id=56149> .

[2] Scarponcini, Paul, *Land and Infrastructure Standards Working Group Charter*, OGC Project Document 12-138r1, October 11, 2012, <https://portal.opengeospatial.org/files/?artifact_id=56138> .

[3] Scarponcini, Paul, *OGC As-Is LandXML-1.2 Conceptual Model*, Preliminary Draft, Version 0.3, June 15, 2013, <https://portal.opengeospatial.org/files/?artifact_id=54380> .

[4] Scarponcini, Paul, *buildingSMART Summit Munich 201310 Meeting Notes*, November 1, 2013, <https://portal.opengeospatial.org/files/?artifact_id=56193> .