Geoscientific
Instrumentation
Methods and
Data Systems

Interactive comment on "The GPlates Geological Information Model and Markup Language" *by* X. Qin et al.

S J D Cox (Referee)

simon.cox@csiro.au

Received and published: 10 July 2012

General Comments ——————-

The paper is primarily a technical introduction to the design of GPML, being an XML format for the serialization and transfer of data associated with the GPlates project and software. There is just a little context given and brief introductions to plate theory, but it is primarily a description of a particular computer application. Hence it is very suitable for this journal.

I only have a general understanding of the Plate Tectonic theory, so will leave it to others to assess that part of the paper. As I have a deeper understanding of the GML tools and the standards context I mainly pick up a few issues in this area.

One general issue about the GPlates and GPML projects that does concern me, and is mentioned briefly in my comments on section 4.8, is that there is no reference to any service, such as a WFS, that publishes GPML data for use in other applications. If this is indeed the case, then how is GPML made available to other applications? How does it practically contribute to interoperability? If there is no GPML publishing service, then the value of defining GPML as a GML Application is questionable. Is it only used internally?

Otherwise, the paper is competent. It will not attract many readers, but it is useful and indeed important to have this material on the record.

Specific Comments ——————

Feature

In section 2.1 it might be worth mentioning that, while a 'GIS Feature' has geographic location, the notion of 'Feature' defined in ISO 19101:2003 "Geographic Information - Reference Model" is more generic than this. Location (a spatial property of features) is not mandatory and may not be applicable.

GML

In section 1. GML is referred to as 'a widely accepted standard'. In fact it is an ISO, AS/NZS and CEN standard which makes it more than merely 'widely accepted'. However, it appears that GPML may be based on GML v2. In which case, why? GML 3 superseded GML 2 and was issued in 2003, and GML 3.2 became an ISO standard in 2007. So any application based on GML 2 is not easily interoperable with GML 3 applications, so a reason for this choice should be provided.

Feature identity

This is introduced suddenly in a very implementation-specific way in 2.1. Identity is a complex topic, with its roots deep in philosophy. While a discussion is clearly out of scope for this paper, the 'endurant' view of (plate) identity is critical to the view

of the world implemented by the technology described in this paper, so some brief introductory remarks or references should be provided.

Feature-type hierarchy

The hierarchy of feature-types is presented in text in section 2.1, in Fig 1 and in detailed lists in Figs 3-5. It might help the reader to provide a tabulation with a text description for each feature-type, as a kind of 'data dictionary' for the model.

Feature property

In 2.2 the authors state "Another important concept in the GPGIM is the feature property". In fact this concept is inherited from GML which implements the OGC Reference Model, which in turn is based on the ISO 19101 Reference Model. Perhaps mention this.

In 2.2.2 a set of types for property values are discussed. The GPGIM types are mostly designed to support discrete representation of functions, parameterized by time or some spatial sampling. This is a general problem, which is addressed in the standard-context adopted by GPML by the concept of the 'coverage' (ISO 19123), which is a spatio-temporal field.

GML again

In the introduction to section 3, GML is introduced as a solution "To alleviate the disorder brought by heterogeneous data". It would be more accurate to say "To alleviate the disorder brought by heterogeneous geospatial data".

Citation for coverages

The ISO/OGC Coverage model is referred to in section 4.5. A citation to ISO 19123:2005 should be provided.

Rasters

Section 4.6 describes Rasters. Again, it should be noted that a Raster is typically a discrete sampling of a spatial function or coverage. Also, netCDF has recently been formalized as an OGC standard (2010) http://www.opengeospatial.org/standards/netcdf .

Interoperability

Section 4.8 reflects on the interoperability implications of GPML. since this is a rather different scope than the rest of section 4, which describes the design of GPML and its implementation as a GML Application Schema, I suggest it deserves its own major section.

However, in this context, a reference to Web Feature Service would be expected. WFS is the standard service for data serialized in GML, indeed a WFS could be characterized as a virtual GML document. The authors are aware of WFS (it is mis-cited in section 3 and 3.1 - see comment below), but do not report providing any WFS service for GPML data.

Figures

Fig 1. What is the reason for the open diamond on the association between InstantaneousFend TimeVariantFeature? It is normally associated with one-direction traversibility towards the end without the diamond. But a common consensus is that the 'aggregation' association is functionally indistinguishable from a non-specific association, particularly if the association is traversible in both directions.

Figs 8-21, 26-29. The XML snippets are both verbose and difficult to read for anyone other than an expert. Consider replacing these with a graphical representation of the schema, such as those provided in most XML IDEs (XML Spy, oXygen).

Technical Corrections —————————

i. In section 1. it is referred to as 'a widely accepted standard'. In fact it is an ISO, AS/NZS and CEN standard which makes it more than merely 'widely accepted'.

ii. the most appropriate reference for GML is

Portele, C. (Ed.)   (2007) Geography Markup Language (GML) Encoding Standard v3.2, OGC Implementation Standard.   OGC document 07-036 http://www.opengis.net/doc/gml (Also published as ISO 19136:2007)

unless it is intended to restrict the GML dependency to GML v2. Note also that the Editors of GML v2 were Cox, Cuthbert, Lake and Martell. The other parties listed were 'contributors', not normally listed in the bibliographic entry.

A number of references for the WFS standard are cited as 'Panagiotis, 2005'. This is wrong: Panagiotis is the editor's given-name, not his family name! - the citation should be to 'Vretanos, 2005'. However, it is also worth noting that WFS v2 was published as ISO 19142:2009.

Interactive comment on Geosci. Instrum. Method. Data Syst. Discuss., 2, 365, 2012.

C65