



Swiss
Seismological
Service

ETH

<QuakeML>

An XML Schema for Seismology

Adrian Wyss¹
Silvio Maraini¹
Danijel Schorlemmer¹
Stefan Wiemer¹
and Manfred Baer¹

¹Swiss Seismological Service,
ETH Zurich, Switzerland.
quakeml@sed.ethz.ch

<Introduction>

Improving access to data and seismic information is a critical issue for seismology. We believe that in order to achieve this goal, the seismological community needs to adopt a flexible, machine independent representation of seismic data.

A major problem related to data exchange in seismology arises from network specific information needs, which thus far has prohibited the creation of one universal, data format. In order to achieve a one-format solution for seismological data interchange, the underlying technique therefore must allow for multi-purpose definitions as well as for user specific extensions, without compromising the format definition or without making the data files unreadable for other users.

We propose an extensible format definition for seismic data (QuakeML) using XML, the eXtensible Markup Language. XML is playing an increasingly important role in the exchange of a variety of data. Numerous business applications, such as stock exchanges, rely on XML as their backbone for data interchange.

Visit us at: www.quakeml.ethz.ch
Or contact us at: quakeml@sed.ethz.ch

See our article in the upcoming
ORFEUS newsletter:
QuakeML - An XML schema for
seismology.
(<http://orieus.knmi.nl/newsletter/>)

<Layer>

QuakeML is divided into several layers. Like a piece of a puzzle, a more sophisticated layer may be added. Layer 2 (Figure 1) is designed to extend layer 1 without modifying any definition already made for layer 1. Layer 1 is designed to store event parameters (see Example). Layer 2 extends this with additional arrival (pick) information. Further layer definitions for other purposes like shakemaps, macroseismic observations, or waveforms may be added.

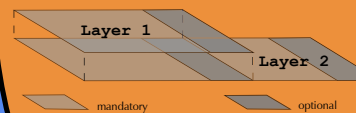


Figure 1: Multi-layer approach: Every layer consists of mandatory and optional data fields. Layer 2 is an extension of layer 1.

<Example>

```
<quakeml>
<event unique_id="EV_01">
<location main="true" unique_id="LOC_01" analysis-type="M">
<origin-date timezone="00:00">
<year>2004</year>
<month>09</month>
<day>28</day>
<hour>17</hour>
<minute>15</minute>
<second>24.0</second>
</origin-date>
<latitude error="0.01">35.8</latitude>
<longitude error="0.01">-120.4</longitude>
<depth unit="km" error="1">7</depth>
<magnitude unit="M" error="0.1">6.0</magnitude>
<region>CENTRAL CALIFORNIA</region>
<author>NEIR</author>
</location>
</event>
```

<Validation>

XML Schema (XSD) is a comprehensive format definition language for describing XML formats. Schemas are supporting rigorous definitions by offering the possibility of defining constraints for every data parameter and of providing metadata like physical units. With the help of such a schema and a parser, an XML-file may be validated during read and write processes. Several parsers already exist.

<Extension>

Any XML-definition can easily integrate additional data. This makes individual extensions of QuakeML possible without compromising the validity. The example (Figure 2) is based on a QuakeML schema; however, the extended data file can be read by any application, which implemented the import according to the original QuakeML schema.

The additional data will simply be ignored. QuakeML is like a puzzle: Just add another piece with you specific parameters.

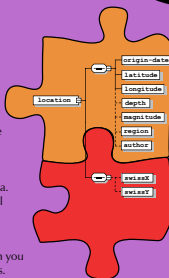


Figure 2: Simplified XML data and a user-specific extension.

<Transformation>

Stylesheet Transformation (XSLT) supports the concept of separating the data from its presentation. XSLT can transform any XML-file into another file format. Several transformation engines already exist. With XSLT, many different presentations of one dataset are possible, e.g., an QuakeML eventlist can become a web page describing the most recent event, an ASCII-list of the main event parameters, or a bulletin of all M>6 events as PDF.

<Summary>

QuakeML based applications offer a number of advantages:

- Import and export routines are readily available.
- Validation and therefore data consistency are guaranteed independently from your code.
- When bringing data to the Internet, the stylesheet transformation is probably the most powerful method available today.
- Adding user-specific data, like application metadata, is possible without making the files unreadable for others.

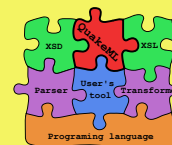


Figure 3: The integration of QuakeML data.

Due to these capabilities, its wide acceptance and the existing large number of utilities and libraries for XML, a structured XML-representation of all sorts of seismological data should in our opinion be developed by fully defining a 'QuakeML' standard.