

XML Wrapper Discussion September 23, 2010

Attendees: Tim Orr and Isaac Holmlund (MN Revisors Office - RO); Dan Dodge (Thomson Reuters); and Robert Horton, Jennifer Jones, Carol Kussmann, Charles Rodgers, and Shawn Rounds (Minnesota Historical Society - MHS)

Purpose: To discuss minor changes to the wrapper schema.

The following account is summarized.

Agenda

- Review values under the Jurisdiction/Body element
- Address any other concerns with metadata set
- Discuss name spaces

4.0 Jurisdiction; Body Element

- Our current acceptable values are House and Senate.
- California tried to apply this to their records; however they use “assembly” which was not an accepted value.
- It was decided that Assembly should be made an acceptable value.

- It was then discussed if the field should include all acceptable values or be free text. It was decided that because there should not be many values to add, it would be better to use a controlled vocabulary in this field.

- California also mentioned that with some of their upcoming documents, they don’t know which body produced the records. This brought up the question of should ‘unknown’ be added as a value? The field is optional but leaving it blank does not necessarily mean that you don’t know who created the records. Adding ‘unknown’ allows you to specify this. It was agreed that “unknown” should be added to the list of accepted values.

- The element will be listed as being extensible if need be.

4.0 Jurisdiction; State

- It was noticed that California used the value CA, California in the Jurisdiction/State field. This value is supposed to be a two character code. (The original MN example also did this.)

- Need to remind CA that the ISO code (3166-2) is the correct value for this field.

- After more discussion it was decided to add a value to the Jurisdiction element. This would allow the state name to become searchable.
 - o 4.0 Jurisdiction
 - 4.1 State Code
 - 4.2 State Name
 - 4.3 Body

3.0 Type

- The current definitions for the document types was deemed to be restrictive. Each state has its own definitions for its terms. Sometimes these terms mean different things in different states. It was decided to remove the definitions.
- In addition to Bill, Resolution, Act, Chapter... add Law, Statute, Rule, and Code.
- Adding more types of documents will help us test if the metadata is extensible across multiple record types.

6.0 Date

- Date block is optional, should this be required? What date is the date referring to?
- Date should be required.
- Required date is publication date. (Change definition to reflect this.)
- Discussion about adding a creation date field in addition to the publication field, but decided against it.
- Date is now: 6.0 Date (R); 6.1 Publication Date (R)

7.0 Session

Make Session Optional (session may not be applicable to all documents)

- Make all sub elements also optional
- Session Year is currently a string value, change this to 'gYear' to get it as a four digit year.
- What is the definition of a session year? The year the document was published, the year within the legislative session? How would California define this? What works for them?

Other Fields

- Type, Body, Governor Action, Event Description, and Rights will be extensible elements.

Name Spaces

- It was found that posting the XML document that California first worked on did not work.
- Dan determined that it was because of the name space, not pointing to a 'real' or existing location
- The value in using a name space is that often you can look at it and determine where the document is from or make an association. It identifies the specific data model for the document.
- The use of `xsi:schemaLocation` instead of `xsi:noNamespaceSchemaLocation` will enable the XML data to better identify which data model, including version of the data model is being used.

Other Discussions

- Talked about connecting documents together and that the goal was really to preserve the artifact, and not the history of the artifact with each wrapper. Information contained within the wrapper however may assist with the history of the artifact.
- Would like to see if we can use the wrapper on different versions of a bill and see if the wrapper can be used to connect the related bills. This would be one validation of the model.
- Another way to track histories or associations would be to add multiple identifiers. The identifier field is repeatable, so in addition to using a unique identifier, you could add the bill number to all the related bills and use this as a way to associate the versions of a bill.

To Do

- Shawn will update the metadata and post it to BaseCamp
- Dan will work on the schema
- Tim and Isaac will update the prototype