XML Registry and Repository Schema Generation

Michael Morris

7/29/10

The next version of the XML Registry and Repository is not totally satisfactory for our purposes. It appears that the sector library will not have a schema creation capability and cannot distinguish between different sector libraries. Since many of our complex types are defined in sector libraries, we would have to include these in the schema manually. To get the support we need from the XML R&R, I think the TAB should clearly define the requirements for sector library schema support. These would include allowing named components to be qualified by namespace identification so that the same named type can be selected across sector libraries. This particular requirement would allow us to include CDS as a separate sector library.

To make the transition easy and to make updating existing standards painless, I propose that the input to the schema creation process would start with a schema file for the particular standard. This file would include the top level element, any new types created for the standard, and the namespace of the target libraries used by schema (sector, CDS, and core libraries). The R&R would then parse the schema, move the new elements to the combined schema, and add type definitions from different libraries as it recursively followed the reference and type attributes of elements and groups. These included types would be unqualified and thus part of the application namespace. The output file would include only the top level namespace declaration and would exclude the import tags. The comments (or R&R processing instructions) with each component would identify the source namespace and version of the component.

Here are two examples of schemas that could be used as input into this process:

ETSR:

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:tsr="urn:org:pesc:message:TestScoreReport:v1.0.0"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:AcRec="urn:org:pesc:sector:AcademicRecord:v1.5.2"
targetNamespace="urn:org:pesc:message:TestScoreReport:v1.0.0"
elementFormDefault="unqualified"
attributeFormDefault="unqualified">
        <xs:import namespace="urn:org:pesc:sector:AcademicRecord:v1.5.2"
        schemaLocation="AcademicRecord_v1.5.2.xsd"/>
        <xs:element name="TestScoreReport" type="AcRec:TestScoreReportType">
                <xs:annotation>
                        <xs:documentation>The complete test score report for multiple
                        students</xs:documentation>
                </xs:annotation>
        </xs:element>
</xs:schema>
```

Recruitment Enrollment:

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns:RecEnr="urn:org:pesc:message:RecruitmentEnrollment:v0.0.2"
xmlns:AdmRec="urn:org:pesc:sector:AdmissionsRecord:v1.0.2"
xmlns:core="urn:org:pesc:core:CoreMain:v1.8.0"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
```

```
targetNamespace="urn:org:pesc:message:RecruitmentEnrollment:v0.0.2"
elementFormDefault="unqualified" attributeFormDefault="unqualified" version="v1.0.0">
        <xs:import namespace="urn:org:pesc:core:CoreMain:v1.8.0"
        schemaLocation="CoreMain_v1.8.0.xsd"/>
        <xs:import namespace="urn:org:pesc:sector:AdmissionsRecord:v1.0.2"
        schemaLocation="AdmissionsRecord_v1.0.2.xsd"/>
        <xs:element name="RecruitmentEnrollment"
        type="RecEnr:RecruitmentEnrollmentType">
                <xs:annotation>
                        <xs:documentation>Top Level Element</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:complexType name="RecruitmentEnrollmentType">
                <xs:annotation>
                <xs:documentation>Major elements for Recruitment and
                Enrollment</xs:documentation>
                </xs:annotation>
                <xs:sequence>
                        <xs:element name="TransmissionData"
                        type="AdmRec:TransmissionDataType"/>
                        <xs:element name="ProspectiveStudent"
                        type="AdmRec:ProspectiveStudentType"/>
                        <xs:element name="SchoolSelectionPreferences"
                        type="AdmRec:SchoolSelectionCriteriaType"/>
                        <xs:element name="NoteMessage" type="core:NoteMessageType"
                        minOccurs="0" maxOccurs="unbounded"/>
                        <xs:element name="UserDefinedExtensions"
                        type="core:UserDefinedExtensionsType" minOccurs="0"/>
                </xs:sequence>
        </xs:complexType>
</xs:schema>
```

The first schema is a released standard and thus all the components would be in the R&R. However, the second schema is still in development and some of the types referenced in the sector library were created specifically for this standard and would not be in the R&R. As a result, any new types and elements not yet approved need to go in the top level schema during development.

Another issue that needs to be addressed is the mechanism for updating to later versions of components. We could use the W3C XML Schema attribute wildcard to tag the version number to a type definition (e.g., <xs:simpleType name="AType" rr:typeVersion="1.0+" rr:typeNamespace= "urn:org:pesc:core:CoreMain">), or we could parse the comments to determine version.

Finally, there is no guarantee that FSA will agree to the requested enhancements and even if they do, the time may be well in the future. To advance this process, I propose that we develop an XSLT script (or other application solution) that will use our current library schemas to create a single namespace schema from its top level definition. During development we would still use the imports but this transform would be applied before release.