Enhanced EEMS XML Schema Validation

This note highlights an issue with the validation of XML documents by the Enhanced EEMS, and proposes a solution for consideration by interested parties.

Background

In anticipation of ongoing UK and EU regulatory changes, one requirement for the Enhanced EEMS is to support changes to the XML schemas used to validate submitted documents. This is especially true for the migration to the Enhanced EEMS at the end of 2010 Q1, when a number of the schemas will change in accordance with the new reporting requirements.

To meet this requirement, the Enhanced EEMS has been designed to maintain a history of each version of a particular schema, and the dates between which that version is valid. For a particular document, it is therefore possible to determine the schema version that applied at the time, and validate accordingly. (For recurring reports this will be according to the date of the end of the report period, and for permit-based reports this will be according to the permit expiry date.)

Example

Consider the DrillFluids.xsd schema, which has significant changes for the enhanced system. Assuming a migration date of 29/03/2010, the system will have the original version of the schema valid to 28/03/2010, and the new version valid from 29/03/2010. This enables any corrections to documents submitted on the original system prior to the migration date to still be validated against the original schema.

Although it is straightforward for the system to choose the correct schema for validation on the server, the current method of exposing the schemas for client-side validation does not handle this scenario. That is, the client would make a request to the EEMS server at:

/eems/OP00/ABC/2010/DrillFluids/DrillFluids.xsd

With just the year being present in the URL, it is not possible for the server to determine whether it should respond with the schema version valid to 28/03/2010, or the one valid from 29/03/2010.

Problem Definition

Although the example given above is in the context of the migration to the Enhanced EEMS, the principle applies to a wider class of changes. However, the issue will only manifest itself for the report types if all of the following conditions are true:

1. The new schema version is valid from any day of the year except 1 January. New versions that start on the first day of a year are not affected.

2. The frequency of the EEMS report type that is validated by the new schema version is bi-annual, quarterly, monthly or permit-based. Report types that are submitted annually are not affected.

3. Documents that were valid for the old schema version are not valid for the new schema version. Minor changes may mean a document validates against both schema versions.

Additionally, these conditions apply not only to the main schemas, but also to any parent schemas that are included by them. For example, if the LibraryUnitOfMeasure.xsd schema changes during the year, all the schemas that include it will be treated as if they have a new version. And for such schemas, it is possible for the year to not be present in the URL, so the server has even less information to use to decide on the correct version to respond with.

Proposed Solution

It is proposed that the server will respond by default with the latest schema applicable to the year present in the URL; in the example above this would be the schema valid from 29/03/2010. For included schemas without a year in the URL, this default would be the schema applicable on the current date.

However, if the client was attempting to validate a submission for earlier in that year, say the correction of a submission for 19/02/2010, then validation would fail. Therefore it is proposed to allow the client to also specify the period end date of the document in the schema URL, so that the server can respond with the correct version:

/eems/OP00/ABC/2010/DrillFluids/DrillFluids.xsd?p_date=2010-02-19

The additional parameter is optional, if not specified then the server's response will be the default described above. In the majority of cases this will be exactly as required.

It is also proposed that for clarity the Enhanced EEMS dashboard will display the URL of the schema required for validation for the minority of documents where the default is not appropriate.

Conclusion

Comments are invited from any interested party in respect of the extended URL format, and any other consequences of this issue that may affect their systems.