NEMSIS Schematron Assessment

Final Report

Prepared by Joshua Legler
June 25, 2018

Introduction

In May–June 2018, I conducted an assessment of the use of Schematron in the NEMSIS standard and in NEMSIS-compliant products, with the goal of recommending changes for NEMSIS 3.5. The assessment included the following activities:

- Evaluate alternative architectures to the **national Schematron schemas** with the goal of making the rule logic and validation messages more readable. Assess performance impacts.
- Evaluate existing **state-level Schematron schemas**. Summarize the rules in the schemas. For each rule summary, include information about the number of states that have implemented the rule, whether the rule applies to national elements, and a recommendation regarding whether the rule should be implemented in the national Schematron schemas.
- Meet with **stakeholders** (which may include state data managers, vendors, and local EMS agency representatives) to gather feedback about the implementation of Schematron in NEMSIS-compliant products.

I delivered a detailed report of findings for each activity. This report is based on those findings and offers recommendations for changes to the NEMSIS Schematron requirements and national Schematron schemas in NEMSIS 3.5 as well as changes to documentation, resources, and technical assistance.

Changes in Schematron Requirements

One stakeholder recommended that NEMSIS compliance certification for "Receive and Process" systems should include the ability to generate Schematron schemas, including logic constructs that must be supported (for example, regular expressions, mathematical operations, set-based comparisons, etc.).

Many stakeholders emphasized the importance of minimizing changes to the NEMSIS Schematron requirements.

Changes in National Schematron Schemas

Changes in the content of the national Schematron schemas would be introduced as part of NEMSIS 3.5. They would not require vendors to make any changes in how they perform Schematron validation; they would simply deploy the schemas. (However, we can expect that changes to the national schemas would trigger analysis and training efforts at the state and local levels.)

Make Constraints More Atomic

Stakeholders expressed support for making constraints more atomic; that is, that there should be a greater number of asserts, with each one covering as few elements as possible, rather than a smaller number of asserts, with each one covering many elements.

Identify Conditional Constraints for Every National Element

For nearly every national element, multiple states have an assertion that conditionally requires the element to have a value. The NEMSIS TAC should consider adding similar assertions to the national Schematron schemas to address every national element.

Few states have assertions that check other aspects of data quality, such as consistency between the responses recorded for various data elements. The national schemas should continue to address such aspects.

Changes in Documentation and Resources

Changes in documentation and resources would provide stronger support to stakeholders as they use Schematron for data validation.

Integrate National Schematron Schema Documentation into the Data Dictionary

Stakeholders expressed strong support for integrating the national Schematron schema documentation into the NEMSIS Data Dictionary. Each data element page should contain a new section that lists the validation rules that involve that data element. The Schematron rules could be incorporated into the data dictionary using an automated approach at the time the data dictionary files are generated. The main benefit of this approach is that it would expose the Schematron validation rules in a clearer way to stakeholders, within the context of each data element's data dictionary page.

Create a Schematron Rule Library

Many states have rules that test the same condition but use slightly different logic or different wording for the assertion text. Several stakeholders suggested that a national library of Schematron rules would help to improve consistency between states. The library could be modeled after the NEMSIS V3 Custom Element Library, and it could include the national Schematron rules as well as state rules, which would help states avoid redundancy or conflict when creating rules.

Provide a Reference Implementation of Schematron Rules to Validate Data Based on StateDataSet, Including Custom Element Data

Many states are interested in deploying rules that enforce constraints that are implied by information contained in StateDataSet documents, such as the following:

- Custom elements
- Certification/licensure levels
- Procedures, medications, and protocols
- Facilities

Some states have implemented constraints on these topics manually, but they have found it difficult to keep those rules current. One state (Oregon) has implemented a more streamlined approach by embedding the StateDataSet in the Schematron schema and keying validation off of it; whenever the StateDataSet is updated, the Schematron schema is simply recompiled, with no changes to Schematron assertion code.

The NEMSIS TAC should publish a reference implementation of a Schematron schema that enforces constraints based on StateDataSet data.

Provide Guidance on Developing, Testing, and Debugging Schematron Schemas

Every assertion in the national Schematron schemas is supported and verified by XML test cases, which demonstrate scenarios where the assertion will pass or fail. Stakeholders reported that state-level Schematron schemas often have bugs. The NEMSIS V3 Schematron Guide recommends that states develop tests cases for their Schematron schemas but does not provide much guidance on how to do so. The NEMSIS TAC should provide additional guidance on developing Schematron test cases for state Schematron schemas.

Develop a Software API Library for NEMSIS Schematron Validation

One stakeholder (a software developer) recommended that the NEMSIS TAC build and publish a set of software libraries for performing NEMSIS-compliant Schematron validation in the most common programming languages/platforms.

Educate States on the Purpose of State Schematron Schemas

Several stakeholders (state data managers as well as software developers) asked for more documentation from the NEMSIS TAC about the purpose of state Schematron schemas and why vendors need them.

Educate States on the Impact of Validation Rules

Stakeholders have varying opinions on the use of errors vs warnings in Schematron schemas, but most recommend leaning toward having more warnings and fewer errors in order to minimize the number of rejected records. Several stakeholders reported a need to educate states on the impact that their validation rules will have on the flow of real-world data.

Clarify What's in the XSD Schema vs What's in the National Schematron Schemas

Some stakeholders reported confusion regarding the scope of the NEMSIS XML Schemas (XSDs) vs. the national Schematron schemas. The NEMSIS TAC should publish documentation that helps stakeholders understand what is covered in XSD vs. Schematron.

Changes in Technical Assistance

Changes in technical assistance are more hands-on and would require additional staff resources at the NEMSIS TAC. Some changes could be partially automated but would involve one-on-one assistance to states as they develop and publish Schematron schemas.

Help States Remove Redundant Rules from State Schematron Schemas

Many states have state-level constraints that are redundant with the national XML Schemas (XSDs). In a NEMSIS-compliant system, these rules will never fire, because the constraints would be enforced during XSD validation. Many states also have state-level constraints that are redundant with the national Schematron schemas. These rules will fire and will result in the user seeing two error/warning messages about the same issue (unless the vendor has made special efforts to deduplicate the messages). The NEMSIS TAC should help states remove these redundancies from state Schematron schemas.

Help States with Versioning and Change Control

States are challenged in managing changes in their Schematron schemas. The NEMSIS TAC should help states to implement changes in their state Schematron schemas that are cognizant of historical data and that account for the time needed by vendors to implement the changes in local systems.

Discourage "Score-based" Rejection of Records in State Systems

The NEMSIS V3 Schematron Guide requires systems to accept records that have no errors and to reject records that have errors (regardless of the number of warnings). Some states accept or reject records based on a validation score derived from the number of failed assertions on a record. This is inconsistent with the NEMSIS V3 Schematron Guide and should be discouraged because it leads to ambiguity regarding whether or not a record is valid.

Review the Content of State Schematron Schemas

This is a potentially time-consuming recommendation, but stakeholders have asked that the NEMSIS TAC review the content of state Schematron schemas for several reasons:

- To identify assertions that are redundant with the national XSD and Schematron schemas
- To identify assertions that have bad logic
- To identify content that is technically valid but not consistent with the NEMSIS standard or recommendations in the NEMSIS V3 Schematron Guide

Other Changes

The following change would need to be implemented in other areas of the NEMSIS standard or have a scope that includes more than just Schematron.

Web Services: Define a Structure for Record-Level Processing Results

In the case of an error, the NEMSIS V3 Schematron Guide allows systems to either reject the record with the error or reject the entire submission. Some stakeholders expressed concern that most systems have implemented the processing of Schematron errors identically to their processing of Schematron fatals, which is to reject the entire submission. It appears that one reason for this is that there is not a standard structure defined in the NEMSIS V3 WSDL that would enable a receiving system list the processing results of a batch submission record-by-record. The NEMSIS V3 WSDL should define a structure for doing so.

Allow Systems to Receive Partial or Invalid Data

Several stakeholders (states) expressed the desire to accept all records into their systems, even if they fail national validation. Among the reasons are the following:

- If a record fails validation at the local level and is never sent, the state doesn't have a way of knowing it. Some states feel that it's better to have everything come in to the state system, then evaluate the data.
- Some stakeholders need to be able to receive partial data (which would not pass NEMSIS validation) for integrations with other systems, such as HIEs.

The NEMSIS TAC should address the issue of how to support some stakeholders' need to accept and process invalid data, while still upholding the standard. One approach may be to modularize the NEMSIS standard so that purpose-built schemas could be used for integrations between systems.

Conclusion

Stakeholders agree that Schematron has been fulfilling its envisioned purpose of enabling data validation rules from the national and state level to be encoded in a standard format and deployed at the local level.

Based on feedback from stakeholders, the recommendations in this report are designed to avoid changes to the NEMSIS Schematron requirements. Instead, most recommendations focus on documentation, resources, and technical assistance to help improve the national Schematron schemas and state Schematron schemas. Changes in the national Schematron schemas would be incorporated into the development of NEMSIS 3.5. Many of the changes in documentation, resources, and technical assistance could be implemented at any time.