Exchange Network Schema Conformance Report Preparation and Review Process

Version: 2.0

Revision Date: June 22, 2006

Prepared for: Network Technology Group

Prepared by: Schema Review Task Force

Document Preparation Process and Acknowledgements

The Network Technology Group (NTG) tasked its Schema Review Task Force with preparing an updated version of the *Schema Review Process for Schema Developers* document in December 2005. The Task Force prepared version 2.0 of this document, which is now titled *Exchange Network Schema Conformance Report Preparation and Review Process*.

The Task Force members and contractors who participated in this effort are listed in the table below.

Task Force Participant	**Affiliation**
Tom Aten	Wisconsin DNR
Dennis Burling	Nebraska DEQ
Charles Freeman	EPA Office of Environmental Information
Nick Mangus	EPA Office of Air and Radiation
Matt Markoff	Ross & Associates
Kurt Rakouskas	ECOS
Bill Rensmith	Windsor Solutions
Nicole Wigder	Ross & Associates

Abstract

This document focuses on the Schema Conformance Report preparation process, which is an integral part of the flow documentation package preparation process. This document begins with an introduction to the flow documentation package preparation guidelines and resources. Next, it details the NTG conformance review process for these flow documentation packages. Last, it provides step-by-step instructions and a template for preparing Schema Conformance Reports.

Table of Contents

Document Preparation Process and Acknowledgements 2
Abstract 2
Table of Contents 3
1. Introduction and Intended Audiences 4
    Purpose of Schema Conformance Report Preparation and Review 4
    Audience for this document 4
2. Schema and Flow Documentation Development 6
    Guidelines for Preparing Exchange Network Flow Documentation Packages 6
    Flow Documentation Preparation as Part of Developing NEW Schemas 6
    Flow Documentation Preparation as Part of Developing UPGRADED Schemas 6
    Tools and Resources for Schema Developers and Flow Development Groups 7
3. NTG Conformance Documentation Review Process 8
    Overview of Conformance Documentation Review Process 8
    What does “Accepting” a Flow Documentation Package Mean? 9
    Accepted Flow Documentation Packages 10
    Flow Documentation Packages That Do Not Sufficiently Meet Network Guidelines 10
4. Step-by-Step Guide for Schema Conformance Report Preparation 12
    A. Confirm that your schema validates and is well-formed 12
    B. Review your schema for compliance with the Exchange Network Design Rules and Conventions (DRCs) 13
    C. Evaluate your schema’s level of integration with the Shared Schema Components (SSCs) 14
    D. Recommendations 15
Appendix: Schema Conformance Report Template 16
    Flow Documentation Status and Contact Information 16
    W3C Conformance and Validation 16
    Design Rules and Conventions Conformance 17
    Shared Schema Components Conformance 17
    Recommendations (optional) 18
    Appendix 1 18

1. Introduction and Intended Audiences

Preparation of a flow documentation package, including XML schema and a Schema Conformance Report, is important for all Network flows. The majority of the flow documentation package is produced to assist future flow implementers. However, flow development groups produce the Schema Conformance Report primarily to assist the NTG with creating and maintaining the standards and guidelines underpinning the Exchange Network (EN).

Section 2 of this document describes preparation of a flow documentation package, Section 3 describes the NTG review process for schema and flow documentation packages, and Section 4 provides step-by-step instructions for preparing a Schema Conformance Report.

Purpose of Schema Conformance Report Preparation and Review

The Schema Conformance Report, and the conformance documentation review process, has been designed to:

- Help flow developers conform with EN standards and guidelines, and improve the quality of schema development on the EN
- Help the NTG collect feedback from flow and schema developers about the relevance of EN standards and guidelines, as well as the effectiveness of EN guidance documents and other supporting materials associated with these standards and guidelines
- Help the NTG better understand implementation challenges, and identify areas where EN governance could enhance its support for flow development

The Schema Conformance Report also provides useful information to schema developers looking for good examples of 'model' EN schema.

Audience for this document

This document is intended to familiarize flow owners, flow development groups, and schema developers with the process for getting schema and flow documentation packages approved for use on the EN. Accordingly, the intended audience for this document is as follows:

- **Flow owner** - the representative(s) of the entity (or entities) sponsoring flow development. Flow owners need to be familiar with the flow documentation package requirements and the conformance documentation review process (Sections 2 & 3).
- **Flow development group** (aka, Integrated Project Team (IPT)) - the group contributing to flow design, and responsible for developing, or managing development of, flow documentation including the Schema Conformance Report. Flow development groups need to understand the flow documentation package requirements, the conformance documentation review process, and the contents of the Schema Conformance Report template (Sections 2 & 3, and the Appendix).

- **Schema developer** - technical resource engaged in developing schema and supporting products/materials under the direction of the flow development group (they may, in some cases, also be a member of the flow development group). The schema developer typically provides major contributions to Schema Conformance Report preparation. Schema developers need to be intimately aware of the EN schema development guidelines and available resources, the conformance documentation review process, the content of Schema Conformance Report template, and the instructions for Schema Conformance Report preparation (Sections 2 - 4, and the Appendix).

2. Schema and Flow Documentation Development

This section provides background information on preparing a flow documentation package, specific instructions for groups preparing new flow documentation or updated flow documentation, and descriptions of tools and resources available to flow development groups.

Guidelines for Preparing Exchange Network Flow Documentation Packages

The Exchange Network *Flow Documentation Checklist*[1] details the required and optional flow documentation that must be included in the flow documentation packages, and links to preparation guidance and templates. ***The Exchange Network governance bodies expect all flow development groups to follow the guidance provided in the Checklist for each Network flow; any questions or issues with the guidance available should be communicated with the NTG[2] as soon as possible.*** Experience has shown it is often easier and more efficient to develop selected flow documentation, such as the Data Exchange Template and the Schema Conformance Report, simultaneously with the flow schemas.

Flow Documentation Preparation as Part of Developing NEW Schemas

Flow development groups preparing new XML schemas should follow all the guidelines listed in the *Flow Documentation Checklist*. The Exchange Network governance bodies strongly recommend schema developers preparing new schemas closely review the Network schema development guidance documents referenced in the checklist and take advantage of the schema development resources listed below.

Flow Documentation Preparation as Part of Developing UPGRADED Schemas

The Exchange Network governance bodies expect that groups upgrading Network schema (major or minor version changes) will update the accompanying flow documentation. Groups preparing upgraded schemas should review all existing flow documentation and determine which documents require updates to align with the upgraded schemas. All upgraded schemas must be accompanied by a Change Control Spreadsheet, which is described on the *Flow Documentation Checklist*. If a Conformance Report was not prepared for previous versions of a schema, a full Conformance Report is required at the next business-driven major version revision. If the flow has an existing Conformance Report, this report does not need to be updated; rather, a Conformance Report Addendum should be prepared, which lists:

- All changes made to the schemas and the reasons for making these changes

[1] The *Flow Documentation Checklist* is available at http://www.exchangenetwork.net/dev_schema/FlowDocChecklist.pdf

[2] Contact information for the NTG is available at http://www.exchangenetwork.net/operations/nob/ntg.htm

- Descriptions of how issues described in the previous Schema Conformance Report or Conformance Report Addendum were addressed, or reasons for not addressing the issues at this time

At the discretion of the flow owner and development group, an entirely new Conformance Report may be generated to accompany a revised schema, if this will be more efficient to prepare or more user-friendly for the reader than an addendum. Flow owners and development groups who are uncertain whether an addendum or full rewrite of the Conformance Report would be more appropriate are encouraged to contact the NTG for guidance.

Tools and Resources for Schema Developers and Flow Development Groups

The Network Operations Board (NOB) and the NTG are committed to providing the necessary tools and support to create high quality Network schema and flow documentation. Specifically:

- The NTG periodically updates schema and flow development guidelines, templates, and tools. Information on these current NTG activities can be accessed on the NTG webpage[3] or during NTG open calls.
- The NOB decided in April 2006 to begin a six month pilot of "early flow development assistance". Under this plan, ***the NOB intends to provide flow development groups with a short contracted technical review of their draft schema structures. This pro-active support is intended to leverage past implementation experience and EN guidance, and to help developers identify and resolve schema (and flow) design issues as early as possible***. This free assistance is optional, but recommended, for new schemas and major schema upgrades. Flow development groups interested in this service should contact the NOB co-chairs[4].

[3] The NTG webpage is available at http://www.exchangenetwork.net/operations/nob/ntg.htm
[4] Contact information for the NOB co-chairs is available at http://www.exchangenetwork.net/operations/nob/membership.htm

3. NTG Conformance Documentation Review Process

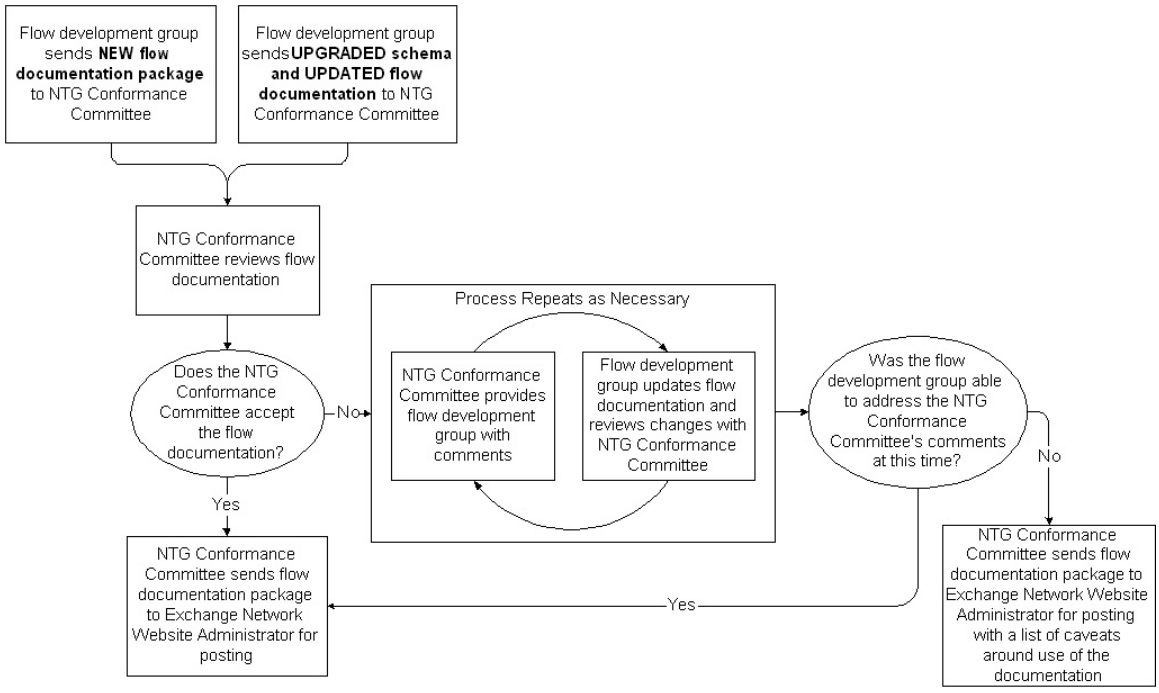
This section describes the process the NTG will follow to review flow documentation packages and post final flow documentation to the Exchange Network website, Registry, and Repository.

Overview of Conformance Documentation Review Process

Flow development groups should submit completed flow documentation packages to the NTG[5], which will form a Conformance Committee (Committee). The Committee will review the documentation and identify whether it meets Exchange Network flow development guidelines. Groups submitting new schemas and flow documentation should submit the entire flow documentation package to the Committee. Groups submitting upgraded schemas only need to submit those schemas, the Conformance Report Addendum, the Change Control Spreadsheet, and any other flow documentation that was updated to align with the upgraded schemas.

The following diagram describes the Committee's conformance documentation review process:

Diagram 1: Conformance Documentation Review Process



The Committee follows the same review process for new flow documentation packages and updated flow documentation. The Committee focuses its review on the Schema

[5] Submit flow documentation packages to the NTG co-chairs. Contact information for the NTG co-chairs is available at http://www.exchangenetwork.net/operations/nob/ntg.htm

Conformance Report or Conformance Report Addendum, and may also choose to review the XML schemas and other flow documentation. After completing its review, the Committee either decides to accept the flow documentation package or decides that modifications are required to the schemas or other flow documentation.

If the Committee decides that modifications are necessary, they will communicate this information with the flow development group. ***It is the responsibility of the flow development group to make the modifications requested and re-submit the complete flow documentation package to the NTG Conformance Committee***. The Committee will review the updated flow documentation package and determine whether further modifications are needed or whether the Committee accepts the package.

What does "Accepting" a Flow Documentation Package Mean?

When the Committee accepts a flow documentation package, it indicates to the Network community that the schema and accompanying flow documentation can be implemented on the Network without caution. Acceptance of a flow documentation package is at the discretion of the Committee, and therefore, it is possible that the Committee will accept flow documentation and schema that do not follow every Network standard and guideline, but still meet the essential requirements for Network flows. For example, this case-by-case acceptance may be applied to schemas meant for both Network and non-Network data exchanges.

If the Committee does not accept a flow documentation package, it indicates on the EN website that the schema and/or flow documentation do not sufficiently meet Network guidelines and standards. This allows the Network community to review the NTG's reasoning for not accepting the flow documentation package to determine whether implementation of the flow is a risk to their agency/department. In cases where a flow documentation package is not accepted by the NTG, EN governance will not prohibit use of the flow. Instead, non-acceptance by the NTG is meant to serve as a caution to implementers of the flow and other schema developers. Members of the Network community are encouraged to bring questions regarding flows whose documentation packages were not accepted to the Conformance Committee.

Accepted Flow Documentation Packages

The NTG Conformance Committee will send accepted flow documentation packages to the Exchange Network Website Administrator (Website Administrator) for posting with the following statement:

> "The [flow name] documentation package has been received and reviewed by the NTG Conformance Committee. The Committee has deemed the schema and supporting flow documentation acceptable for use on the Exchange Network."

The Website Administrator will use the following process for posting the flow documentation package and statement listed above:

- Exchange Network Repository: Website Administrator will post the schemas to the Repository
- Exchange Network Registry: Website Administrator will send the schemas and other flow documentation to the EN Registry Administrator (currently at EPA)
- Exchangenetwork.net site: Website Administrator will post the statement listed above next to the flow documentation on the website.

Flow Documentation Packages That Do Not Sufficiently Meet Network Guidelines

The NTG expects that schema development groups will work with the Committee to fix any identified points not meeting Network guidelines. However, the NTG understands that occasionally flow development groups may be restricted by resources or time, and therefore may not be able to complete the necessary modifications to their schemas and flow documentation. In these cases, the Committee will still forward the flow documentation package to the Website Administrator for posting, but will take steps to inform other Network participants of the identified points not meeting Network guidelines.

First, the Committee will prepare the following statement:

> "The [flow name] documentation package has been received and reviewed by the NTG Conformance Committee. The Committee determined that ***the schema and flow documentation package does not sufficiently meet the Exchange Network flow development requirements. Therefore, Network participants should implement the flow with caution. The following requirements were not met:***
>
> 1. [name and/or description of first requirement not met]
> 2. [name and/or description of second requirement not met]
> 3. Etc.
>
> Please refer to Conformance Report for further details regarding the compliance of this schema with Exchange Network guidelines."

This list will include specific schema or flow development guidelines that were not met, including:

- Improper use of schema development guidelines (e.g., non-use of the SSCs or namespace formatting)
- Failure to provide required documentation (e.g., the Data Exchange Template)
- Other issues at the discretion of the NTG Conformance Committee

Using the statement above, the Conformance Committee may take the following steps in conjunction with posting the schema and flow documentation:

- Include the statement along with the flow documentation on the exchangenetwork.net website.
- Include a text file with the statement along with the schema in the Exchange Network schema repository.
- Include a text file with the statement along with the zip file in the Exchange Network registry.
- Require that the status of the schema and flow documentation be noted at the top of every document associated with the flow (e.g., Flow Configuration Document, Schema Users' Guide).
- Request that the flow development group re-version the schema to less than a 1.0 version (for new schemas only).

4. Step-by-Step Guide for Schema Conformance Report Preparation

This section of the document provides a step-by-step guide for schema developers to prepare a Schema Conformance Report. A Schema Conformance Report template is available in the Appendix.

A. Confirm that your schema validates and is well-formed

Under this step, flow development groups should complete two tasks: confirming the schema files are well-formed, and confirming the schema files and instance documents validate using all XML parsers that will be used in conjunction with the flow.

Review your schema for W3C compliance using W3C XML Schema Validator (XSV).

1) Place your schema files on a web server accessible to the internet.
2) Generate a list of the full URLs to each of your schema files. This can be done in many ways; one method is to get a directory listing for the directory on your web server where you have placed your schema files. After copying and pasting this listing into Excel, use the 'Text to Columns' function to extract the filenames of each of your schema. Then, using worksheet functions, construct a full URL for each filename, and link them together into a single space-separated string.
3) Open the W3C's XSV tool at http://www.w3.org/2001/03/webdata/xsv in a web browser. Enter the full URL to each of your schema files into the 'Address(es)' text box. Click the 'Show Warnings' and 'Check as complete schema' checkboxes. Click the 'Get Results' button.
4) If the 'Get Results' button is not responsive, you may need to do the validation of your schema files in batches (there is a limit for the input text box of around 1600 characters).
5) Fix any errors identified in the results and repeat Steps 1-5 until the tool finds no errors (note that you do not need to address the warnings identified by the tool).
6) Copy and paste any warnings identified in the results into the table provided in the Schema Conformance Report template and provide a brief explanation of why the warning occurred. If the same warning is repeated many times, a single explanation will suffice.
7) Copy and paste the entire output of the XSV tool into the appropriate section at the end of the Schema Conformance Report.

Validate your schema and example instance documents using all XML parsers that will be used in conjunction with the flow. Work with all flow partners to make sure that the schema and instance documents validate using all parsers that will be employed in the actual exchange. There are many tools[6] that may be useful for validating schema and instance documents during development. For the purposes of the Conformance Report, you are asked to determine which parsers will be used in conjunction with the

[6] Such as: XMLSpy from Altova, IBM Schema Quality Checker (SQC), Xerces2 Java Parser, Sun Microsystems Multi-schema XML Validator, and the Microsoft XML parser.

flow and, at a minimum, ensure that your schema and instance documents validate using these parsers.

Since EPA's Central Data Exchange (CDX) is involved in many EN flows, validating your sample instance documents using the CDX parser is also required. CDX hosts a validation service that is available to all EN participants. It can be called as web service by any EN node, or called via a web browser interface. This service is available at https://tools.epacdxnode.net/. Before using this service to validate instance documents against a new or revised schema, you will need to work with CDX to have your schema files loaded in this service. Contact the CDX help desk at nodehelpdesk@csc.com.

B. Review your schema for compliance with the Exchange Network Design Rules and Conventions (DRCs)

There are two stages in reviewing your schema for compliance with the DRCs. The first stage uses the Schema Design Tool (aka, schema checker tool), and the second stage is a manual process to ensure compliance with the DRC v1.1 (which is slightly more up-to-date than the Schema Design Tool).

Use the Schema Design Tool to check each of your schema files.

1) The Schema Design Tool is available from the Facility Registry System (FRS) web site at http://www.epa.gov/enviro/html/frs_demo/design_tool.html.
2) Enter the full path and file name of the XML schema. Click the "submit" command button. The output results will be displayed on a web page.
3) By copying and pasting into another document, or by using the browser's standard print or save commands, retain a permanent copy of the results for each file.
4) After you have processed all of your schema files through the Schema Design Tool, you may wish to produce a consolidated version of the tool's output so that you have a list of all the DRC violations for the next step. Do not enter any of the output into the Schema Conformance Report at this time.

Review your schema and the schema checker output against the DRC v1.1[7] to confirm compliance with the most up-to-date Exchange Network guidance. Appendix A of the DRC v1.1, which contains the complete list of updated XML design rules and conventions, is suggested as a reference for the steps below.

1) Manually inspect each violation identified by the Schema Design Tool against the DRC v1.1. Any violations identified by the Schema Design Tool that are not in violation of the DRC v1.1 do not need to be listed in the Schema Conformance Report.
2) In the table provided in the Schema Conformance Report template, list the DRC violations identified by the Schema Design Tool that were not eliminated in Step 1 above. For each violation, include the DRC Rule ID (e.g., SB-4), the filenames of the schema where the violation occurs, and an explanation of why the rule was

[7] http://www.exchangenetwork.net/dev_schema/EN_XML_Schema_Design_Rules_v1.1e.pdf

violated. You must include all cases where a 'SHOULD' or 'MUST' rule was violated in this table. Reporting on the use of 'MAY' or 'SHALL' rules is optional, but strongly recommended in cases where the developer's explanation provides feedback regarding the rule.

3) Conduct a manual review of the DRC v1.1 and your schema files to identify any violations that were not identified by the Schema Design Tool. List any further violations identified in the Schema Conformance Report table, following the same instructions provided in Step 2. A number of important changes to EN guidelines are contained in the DRC v1.1. Special care is recommended in assessing compliance with the design rules and conventions around namespace formatting and declaration, file naming, and versioning.

C. Evaluate your schema's level of integration with the Shared Schema Components (SSCs)

This part of the review focuses on compliance with Shared Schema Components (SSCs) and providing feedback that may help improve the SSCs in the future.. The first stage involves documenting the level of integration with the SSCs. The second stage documents any cases where possible SSC integration was rejected. The third stage identifies content areas where future SSCs could be developed.

Document your integration of the SSCs

1) Document in the Schema Conformance Report all instances where your schema makes use of the SSCs. Each instance should be categorized as High, Medium, or Low integration with the SSCs, as defined below. The Schema Conformance Report template provides three tables for documenting each level of SSC integration.

    - High Integration: instances where SSC elements or data types with complex content were directly integrated into the target schema without modification. For each instance of high integration, list the element name and the schema path to that element in your schema.

    - Medium Integration: instances where SSC elements or data types with complex content were modified through the process of XML extension and/or restriction before being included into the target schema. For each instance of medium integration, list the element name, the complex SSC element or data type that has been modified, and whether the derivation was by extension or by restriction.

    - Low integration: instances where elements or data types with simple content were integrated into the target schema. For each instance of low integration, list the element name and the simple SSC element or data type.

Document any cases where you decided not to use available SSCs

1) Working from your DET, review each element in your schema to check for cases where an SSC element or data type might have been used. For each element compare against the SSCs (using either the SSCs themselves[8], or the SSC Usage Guide and Technical Reference[9]) to identify possible integration opportunities. Record these opportunities in the Schema Conformance Report.
2) If the SSC element or data type is easy to integrate in any of these cases, you are expected to make a good faith effort to modify your schema to use the available SSC.
3) If you integrate any further SSCs as a result of this analysis, add this information to the High, Medium, or Low integration tables in the Schema Conformance Report.
4) If you are not able to, or choose not to take advantage of these integration opportunities, you must document your reasoning in the Schema Conformance Report. For each instance where an SSC element or data type could have been used, document the following information in the Schema Conformance Report table: the name of the source schema, the element or type name, the corresponding SSC element or data type, and an explanation of why the SSC was not integrated.

Identify areas for future SSC development

1) Having just reviewed, in the step above, every element in your schema and the SSCs currently available, please now help identify candidates for future SSC development. In addition to suggesting elements and data types that seem to be natural candidates for reuse on the EN, please also consult the CRM 2.0[10] for large data blocks that are already part of the EN data model[11]. When documenting candidates for new SSCs in the Conformance Report, please note any corresponding CRM data blocks.

D. Recommendations

In the final section of the Schema Conformance Report, you may optionally document any comments or recommendations you have for the Exchange Network governance bodies based on your experience preparing the Schema Conformance Report and flow documentation package. For example, these recommendations could pertain to the Step-by-Step Guide for Schema Conformance Report Preparation, Exchange Network schema development guidelines, or other Exchange Network supporting materials.

[8] Available at http://www.exchangenetwork.net/schema/SC/
[9] Available at http://www.exchangenetwork.net/dev_schema/
[10] Available at http://www.exchangenetwork.net/dev_schema/CRMv2_0.doc
[11] Note that SSCs currently exist for some CRM 2.0 data blocks, but not all.

Appendix: Schema Conformance Report Template

Flow Documentation Status and Contact Information

Flow Name:

Schemas/Versions included in Conformance Report:

Conformance Report Author:
Contact Information:

Schema Developer:
Contact Information:

Flow Owner or Other Point of Contact for Flow Documentation Package:
Contact Information:

Date Flow Documentation Package Submitted:

W3C Conformance and Validation

W3C's XSV Tool Output:
[ ] yes [ ] no Output pasted in the last section of the Conformance Report
[ ] yes [ ] no Zero errors identified by XSV Tool

Explanation of Warnings Listed in XSV Tool Output:

Warning	Explanation

Schema and Instance Document Validation:
Names of XML parser(s) intended for use in conjunction with the flow and development tool(s) used to validate the schema and instance documents:

[ ] yes [ ] no All schema files validate using all parsers and tools listed above
[ ] yes [ ] no All sample instance documents validate using all parsers and tools listed above
[ ] yes [ ] no All sample instance documents validate using the CDX validator service

Design Rules and Conventions Conformance

Explanation of DRC Violations Identified:

DRC Rule ID	Schema Filenames	Explanation

Shared Schema Components Conformance

High Level of SSC Integration:

Element Name	Schema Path

Medium Level of SSC Integration:

Element Name	Complex SSC Element or Data Type	Derived By

Low Level of SSC Integration:

Element Name	Simple SSC Element or Data Type

Explanation for Not Integrating Available SSCs:

Schema Name	Element or Type	SSC Element or Type	Explanation

Candidates for New SSCs:

Schema Name	Element or Type Name	CRM data block

Recommendations (optional)

1.
2.
3.

W3C's XSV Tool Output

<insert output from the XSV tool here>