1. Title: Schema Review Process for Schema Developers

2. Author: TRG/Schema Review Work Group

3. Subject: The process proposed for schema developers or sponsors to take in order to document a schema's compliance with Exchange Network requirements at Recommendation Track milestones.

4. Description: This document proposes that each Network schema be evaluated for compliance with the XML Design Rules and Conventions, with the modularity aspects of the Core Reference Model, and with the data element standards established for Network use. The evaluation may be done by the schema developer or Sponsor, or by an independent party. This document also describes the documentation package necessary for TRG review in order to establish the Recommendation Track status of the schema.

5. Publisher: TRG

6. Contributors: Schema Review Work Group:

Tom Aten – WI	Brand Neimann – US EPA
Bruce Bargmeyer - US EPA	Dennis Burling – NE
Jeff Kohn – US EPA	Larry Fitzwater – US EPA
Louis Sweeney – Ross & Assoc	Mike Beaulac – MI
Molly O'Neill – ECOS/NSB	Sarah Calvillo – Ross & Assoc
Steve Vineski – US EPA	Terry Forest – US EPA
Tim Crawford – US EPA	

7. Date: This document is released at the Last Call Working Draft on December 1, 2003.

8. Comments: Comments on this document should be sent to Tom Aten, thomas.aten@dnr.state.wi.us

9. Comment Period: December 1, 2003 to January 2, 2004.

10. Forum: This document and any comments received during the comment period will be posted on the Exchange Network web site (www.exchangenetwork.net).

11. Type: This is a Report of Significance;

12. Source: Schema Analysis Results and Recommendations Report, service agreement: NE-TRG-02, prepared for ECOS/NSB by SAIC.

13. Changes: The source report was changed to apply to schema developers or their sponsors.

14. Coverage: The process and documentation package described by this document is intended to apply to all schema developed for use within the Exchange Network.

15. Rights: This document is publicly available for reuse and distribution. Neither the TRG, NSB, EPA, nor the Exchange Network claim any rights to the content of this document.

16. Support: Members of the Schema Review Work Group may be contacted for additional information.

17. Implementation: Implementation experience is not required at this time, but any such application would be appreciated.

18. Data Flow: Not Applicable.

19. Endorsement: Last Call Working Draft

20. Stability: The process and documentation aspects described by this document are expected to remain stable. This document references resources outside of this document, in particular the XML Design rules and Conventions, the Core Reference Model, and Data Standards. These are also in the development process. It is anticipated that these will become mature in 2004.

A SCHEMA REVIEW PROCESS FOR SCHEMA DEVELOPERS

PURPOSE:

This document describes the package of materials, the resources available, and the steps a schema reviewer needs to take to review a schema prior to submitting it to the Technical Resource Group (TRG) for recommended use by Exchange Network participants. The TRG has established an Approval Process, which includes increasing levels of maturity, to provide peer review of schema for Exchange Network use.

INTRODUCTION

The Schema Review Process is an important step in getting your schema ready for exchanges on the Environmental Information Exchange Network (a.k.a., Exchange Network (EN)). The Review Process ensures that schema adhere to the values and standards of the Exchange Network. There are many ways to find out more about the Exchange Network, including:

- Exchange Network Grant Guidance Documentation
- Exchange Network Website (http://www.exchangenetwork.net)
- Grant award documents

These sources will provide you with the goals of the EN and advise you to initiate your schema development effort based on the eXtensible Markup Language (XML) Design Rules and Conventions (DRC), existing schema, documented elements from the Environmental Data Registry (EDR), Environmental Data Standards Council (EDSC) data standards, and modularity defined by the Core Reference Model (CRM). These are critical for schema designed to flow data within the EN, including facility to state, state to state, state to Environmental Protection Agency (EPA), and all combinations in between. To help you meet these goals, the Technical Resource Group (TRG) developed this process to promote your schema for use by EN participants.

BACKGROUND:

The EN is a major EPA/State initiative to promote data exchange and information sharing across various environmental programs and among EN Partners (Partners). A major tool of the EN is the XML language and the development of XML schema to be used to exchange environmental data among Partners. A number of efforts are underway to ensure the success of the EN, including the Environmental Data Standards Development, the CRM, XML DRC, this Schema Review Process, and the TRG Approval Process. These efforts are envisioned to aid schema developers by reducing redundancy and encouraging harmonization among schema.

In the summer of 2003, the TRG conducted a Schema Review Pilot Project that recommended a process for technical reviews of schema. These recommendations are incorporated into this document to guide schema developers through a technical review that evaluates schema for the:

- Compliance with World Wide Web Consortium (W3C) XML schema requirements and standards (e.g., validates, well-formed, etc.)
- Conformance to the EN XML DRCs
- Comparison to CRM for overall modularity and identification of potential new data blocks that could be added to the CRM
- Conformance with applicable data standards (http://www.epa.gov/edsc/)

Appendix A is a description of the steps involved in the TRG Approval Process and a diagram illustrating the relationship between the schema review and the approval process.

SCHEMA REVIEW:

Schema Review is the technical review process to prepare a schema to be approved/certified for EN use by the TRG. For this review process, "schema" refers to the elements, tags, structure, modules, etc., that make up a proposed data flow.

The following is an outline of the steps to be taken in the schema review:

A. Schema sponsor/developer determines initial review requirements and responsibilities
B. Conduct Schema Review
    a. Review schema for compliance with W3C schema requirements and standards
    b. Review schema for conformance with EN XML DRCs
    c. Compare schema with the CRM
    d. Review schema for conformance to Environmental Data Standards
C. Prepare Schema Review Conformance Report
D. Deliver Schema Review Submission Package

Schema Review Process:

<u>Step A: Determine Reviewer and Supporting Materials</u>
The TRG is proposing that either schema developers or an independent party will perform the XML schema review prior to submitting a schema to the TRG for consideration in the Approval Process. The program/agency sponsoring the XML schema is most familiar with the schema's content and is perhaps best suited to make the comparisons at the data element level. If the sponsoring program/agency elects to choose an independent party to perform the review, then the sponsoring program/agency and the schema reviewer should determine the scope of the review, identify any necessary materials they will require, and develop a detailed schedule for the review. The following materials should be used for the review process:

- Electronic copy of the complete XML schema package under review.

- Data dictionary describing the content of the schema (including element names, definitions, sizes, and formats). This information may also be imbedded in the schema.
- Code sets with meanings including the locations of code sets incorporated by reference.
- Trading partner agreement describing the information being exchanged (if available).
- Entity-relationship diagram (ERD) that indicates the reporting requirement (mandatory, required, optional) of each data element being exchanged (if available).
- XML schema user's guide (if available).

The following templates are meant to aid reviewers of XML schema. Other helpful information such as the DRCs, CRM, etc. can be found on the EN website (www.exchangenetwork.net).

- Schema Design Rules and Conventions Checklist Template (Appendix B)
- CRM Harmonization and Modularity Checklist (Appendix C)
- CRM Review Matrix (Appendix …)
- EDSC Element Comparison (Appendix D)
- Summary Results Matrix template (Appendix E)

Note for Last Call Working Draft period: The appendices listed above, and others meant to support this review process, will be finalized during the LCWD review period. This document with the appendices will be considered for the Candidate Recommendation Status Decision.

Note for Last Call Working Draft period: The CRM and Data Standards may produce a combined analysis tool through the Architecture work that would replace the three appendices listed above and referenced in Step B.3 and B.4.

Step B: Conduct the Review

To accomplish the review electronically, the following tools and resources are needed:

- Spreadsheet or word processing software (e.g., Microsoft Excel)
- Schema validation software (e.g., XML Spy)
- Database software (e.g., Microsoft Access 2000) to import data from XML Spy
- The Schema Design Tool (www.epa.gov/enviro/html/frs_demo/new_docs.html)
- EDSC data standards (www.epa.gov/edsc/)
- An electronic copy of the CRM (www.exchangenetwork.net)
- This process document and data reporting templates provided in the Schema Review Package (www.exchangenetwork.net)

Step B.1: Review Schema for Compliance with W3C Schema Requirements and Standards

A number of tools can be used to validate a schema against the W3C standards, for example XML Spy. You must ensure compliance with the W3C standards as the first step in the review process. If a schema does not validate, no further review can be done until the XML schema is successfully validated. Here is an example using XML spy:

1. Create project in software (XML Spy) that validates XML schema against W3C standards. This project should include all the XML schema that will be reviewed.

2. In the project window, highlight the name of the project.

3. Click the Validate icon (green check) on the XML Spy toolbar.

4. If all schema have validated successfully, a message box will be displayed saying that all schema have been validated successfully.

5. If a schema has an error, remove that schema from the project and return to step 3.

Step B.2: Review Schema for Conformance with the Exchange Network XML Design Rules and Conventions

Review each XML schema for conformance to the design rules. Use the Schema Conformance Review Matrix to identify the evaluation criteria for conformance to the design rules. The results of the Schema Design Tool, XML schema validation, and manual inspection are the mechanisms used to determine compliance. Complete the Detail Results Matrix for each XML schema evaluated. The Schema Design Rules Checklist Template can be used to develop the Detail Results Matrix. The Template is designed to hold six XML schema modules, if the application has more than six schema modules prepare multiple matrices.

1. Process all XML schema through the Schema Design Tool. 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. 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. Use the browser's standard print or save command to retain a permanent copy of the results.

2. Manually inspect the schema where the criteria indicate.

3. Develop a Summary Results Matrix for the application. This matrix summarizes the findings. It is used to provide overall counts and enter general comments. A Summary Results Matrix template (see Appendix E) is provided.

4. Prepare response report, using the detail and summary matrices, and your justifications of deviation from the standards. Common deviations, for example those that recur in the schema, should be discussed only once.

Step B.3: Compare the XML Schema with the CRM

For each XML schema prepare a list of elements and attributes with tag name and data type/size information. This list will be used for both the CRM and data standards review. This process may be done by manual review of the schema or by extracting the elements into a database (e.g., Access 2000) using a tool, such as the Export to database function of XML Spy. The list should contain the name of the source schema, the element and attribute XML tag, data type and size, and annotation and definition.

1. Use the element and attribute list and other documentation, such as the data dictionary or annotation comments, to prepare for comparing elements to the CRM data groups and blocks. Sort the list in alphabetical order and add a column for the number of the matching block and a CRM review comment column.

2. Using the name of the element or attribute XML tag, search the CRM document either electronically or manually to see if it matches any of the information reviewed in the document. If there is a match, record the block number on the worksheet. If there is not a match, manually review the CRM for information blocks related to the element. If there is still not a match, note that the element may be a candidate for a new data block.

3. Complete the CRM review matrix including suggestions, modifications, and comments.

Step B.4: Review the XML Schema for Common Data and Data Standards Conformance

1. Using the B.3 list of element and attribute XML tags and data types/sizes, compare it to the XML tags and sizes in the EDSC data standards. A notation should be made in the comment column when data covered by a data standard is reported using non-standard XML tags or field sizes and when mandatory information is missing.

2. If the review is to include the enumerated values associated with the schema, review the schema manually to determine which elements have value lists. Compare the lists with standard values.

3. Prepare a data standards report including suggested modifications, additions, and comments.

Step C: Prepare Schema Review Conformance Report

Use the templates and any additional documentation from each step of the review to prepare a draft report. The concise draft review report should include the following sections:

- Introduction: Describes the background and purpose of the system reviewed.

- Methodology: Provides an overview of the analysis portion of the review process, including the standards and business rules examined, and version indications.

- Findings: Summarizes the results of the review. This section should clearly indicate apparent discrepancies/deviations and present your justification.

- Recommendations: Identify issues with the review materials or XML design concepts that the TRG should consider in advancing schema use in the EN.

- Supporting Appendices: Appendices should include the templates, checklists/matrices, and Schema Review Tool outputs developed during the review.

If an independent party performs the schema review, the sponsoring program/agency should hold a meeting/teleconference once a draft report is complete to discuss the findings of the review with the reviewer (independent party). The conformity review detail supporting documentation should be carefully reviewed to determine whether the analysis accurately reflects the XML schema and its documentation. Once inaccuracies in the review have been addressed, the discussions should focus on bringing the XML schema into conformance. Any comments provided by the review should be incorporated in the final report.

Step D: Schema Review Submission Package

Once the final Schema Review Report is complete, a schema developer/sponsoring program/agency should submit the review report package to the TRG to initiate the schema review approval process. To submit, schema sponsors/developers representatives will fill out an online application found on the Registry website (www.exchangenetwork.net, click on Registry (to be developed winter '04)). The review report package should include the items listed in Step C, and a section about the status of the document. The status section should provide the following:

21. Title: A name given to the resource;

22. Author: An entity primarily responsible for making the content of the resource;

23. Subject: A topic of the content of the resource;

24. Description: An account of the content of the resource;

25. Publisher: An entity responsible for making the resource available;

26. Contributors: Entities responsible for making contributions to the content of the resource;

27. Date: A date of an event in the lifecycle of the resource, typically associated with creation or availability;

28. Comments: Where to send comments and comment period;

29. Comment Period: Dates when comments will be accepted.

30. Forum: Where the document and comments are published;

31. Type: The type of document it is (i.e., Report or Technical Recommendation);

32. Source: A reference to a source from which the present resource is derived;

33. Changes: Significant changes from the previous version;

34. Coverage: The extent or scope of the content of the resource;

35. Rights: Information about rights held in and over the resource;

36. Support: Designated support contacts or mechanism, if any;

37. Implementation: Whether implementation experience is being sought;

38. Data Flow: For XML Schema, information about the data flow;

39. Endorsement: Level of TRG Support for the recommendation; and

40. Stability: Expected amount of change as it progresses through the Recommendation Track.

APPENDIX A: TRG SCHEMA APPROVAL PROCESS:

The TRG Approval Track is the process the TRG will undertake to approve/certify a schema for EN use. This process is incorporated into the broader TRG review and recommendation track for technical documents. For schema, a review conformity report must be submitted for a schema to change status (e.g., Last Call Working Draft to Candidate Recommendations). The following review process is based on section 4 of the TRG Review Process for Technical Reports and Recommendations document.

Steps for the TRG Approval Process are to

A. Review the submitted *Schema Review Submission Package* to fulfill requirements for status determination
B. Determination of package completeness
   a. Package complete
   b. Package complete with requested deviation from review guidance
        i. Includes the documentation (proposed defense) to support the deviation
   c. Package incomplete
        i. TRG returns and requests additional documentation
C. Approval Track decision
   a. Accepted, fulfills requirements to move to next status level
   b. Provisionally accepted
   c. Not Accepted
        i. Returned for additional work and documentation
             1. Left at the current status level
       ii. Does not meet requirements or requested deviations are not acceptable to move to the next status level, will remain at the current level with no further modification proposed
      iii. Schema development effort is split, this action is held at current level, a new effort is (required?) to be initiated to develop a new compliant schema!

The following diagram outlines the Schema Review and Approval Process

TRG Schema Review and Approval Process

Step 1: Schema Review	Step 2: TRG Approval Track

Step 1:
- A: Schema sponsor or developer determines review
- B: Conduct schema review
- C. Prepare Schema Review Conformance Report
- D. Submit Schema Review Submission Package

Step 2:
- A. TRG review of Schema Review Submission Package
- B. Determine completeness of package
  - Complete
  - Request deviation and proposed defense documentation
  - Incomplete
- C. Approval Track Decision
  - Accepted → Status changed
  - Provisionally accepted
  - Not Accepted → Frozen or returned