Electronic Commerce Connection, Inc.

Publication/Content Management and the NDR

Prepared for:

Betty Harvey

harvey@eccnet.com

Publication/Content Management and the NDR

Betty Harvey

Electronic Commerce Connection, Inc.

harvey@eccnet.com

Table of Contents

Introduction

Section 1: Background

Slide 1: Publication/Content Management and the Navy NDR

Betty Harvey
Electronic Commerce Connection, Inc.
harvey@eccnet.com

Slide 2: Background

Currently aiding NIST in developing validation methodologies for Naming and Design Rules. (NDR)

- Initial work using Schematron (ISO/IEC 19757) to validate Navy NDR's.
- Total of 129 rules.
- Some rules can be broken into multiple rules.
- Rules are in following functional areas:
    - Attribute Declaration
    - Code List
    - ComplexType Definition
    - ComplexType Naming
    - Documentation
    - Element Declaration
    - Element Naming
    - General Naming
    - General XSD
    - Instance Document
    - Information Analysis
    - Modeling Contraints
    - Namesapce
    - Root Element Declaration
    - Security
    - Schema Structure Modularity
    - Standards Adherence
    - Versioning

Slide 3: Implementation Rule

Estimated percentages of NDR rules that can be validated using Schematron.

Schematron Rules	Estimated Percentage
Testable	30%
Partially Testable Rules	35%
Untestable	35%

Slide 4: EDI's Relationship to NDR

- EDI XML Initiative start ed in late 1990's.
- Guidelines for using XML for Electronic Data Interchange, January 1998
- ebXML family of specifications was a bridge between EDI and XML
- ebXML - approved May 2001
- UBL - Defines schemas for EDI.

Jon Bosak: *"The UBL technical committee started when some ebXML participants wanted to define schemas for XML payload that is exchanged"*

NOTE: EDI - Electronic Data Interchange

Slide 5: EDI vs Documents

- EDI and Document standards developed independently
    - EDI - X12/ISO 9735 EDIFACT
    - Documents - ISO 8879 SGML/ISO/IEC 8613-1:1994 ODA
- EDI people didn't communicate with Document people
- Document people didn't communicate with EDI people
- The communities didn't feel like they had anything in common with one another

XML Brought The Groups Together!

Slide 6: Traditional EDI

- Computer-to-computer exchange of business data.
- Hands-off exchange of data between computers.
- Agreement between 2 parties
- Terse messaging system
- Paper-based standards
- Usually batch transactions

Slide 7: Paper-based EDI Standard

Sample EDI Standard

Slide 8: EDI Message

- Concise
- Delimited
- Implicit Mapping
- Segment Based

X12 Name Segment Example

```
N1*SH*ACE MANUFACTURING*1*987654321*
N2*RECEIVING*N3*234 MARKET STREET*N4*
SAN FRANCISCO*CA*94103*US
```

Semantic Meaning

N1	EDI Semantics
SH	Shipper
ACE MANUFACTURING	Name
1	Type of Identifier
987654321	DUNS Number
N2	Additional Name Information – Optional
RECEIVING	Division
N3	Address Information
234 MARKET STREET	Street
N4	Geographic Information
SAN FRANCISCO	City
CA	State
94103	Postal Code or Zip
US	Country

The example above is delimited with astarisks (*).

Slide 9: EDI/XML Commonality

EDI	XML
Need defined rules	Need defined rules
Structure is implicit	Structure is explicit
Machine => Machine	Maching => Machine
	Machine <==> Human
Widely used in large industries	Widely used in large industries
Expensive for SMEs*	Cost effective for SMEs*
Conversion capable for WWW	WWW enabled

   * Small and medium-sized enterprises

Section 2: Documents

Slide 10: What is an Electronic Document

Websters: *an original or official paper relied on as the basis, proof, or support of something c : something (as a photograph or a recording) that serves as evidence or proof 2 a : a writing conveying information b : a material substance (as a coin or stone) having on it a representation of thoughts by means of some conventional mark or symbol*

Exchange of information:

- Computer processable information
- Human processable information
- Information that can be manipulated to create other types of information.

Slide 11: Government Information Flow

Federal government is unique. It has both vertical and horizontal business requirements.

- Internal Agency
- Agency-to-Agency
- Supplier-to-Government*
- Government-to-Supplier

NOTE: *Government vendors also supply to commercial entities. They may have business requirements to supply information in both commercial and government standards - airplane manufacturers are a good example.*

Slide 12: Government Specifications/Requirements

- Ambiguous requirements increase cost to the government suppliers.
- The more complex a specification, the more the product or service costs.
- Costs are transferred from supplier to government and ultimately the taxpayer
- Specification live in government longer than they are practical to maintain
- Most government suppliers also supply goods and services to commercial clients. They support multiple standards.

NOTE: In many cases obsolete technology tends to be required in RFP's longer than they are supported. Guidance doesn't trickle down the organization(s) fast enough.

Section 3: Facilitation
of Documents in NDR

Slide 13: Facilitation of Documents in NDR

These next few slides will address areas of concern for developing schemas for documents and in the exchange of document-type data while complying with the Navy NDR.

I will be using a schema that was developed for accreditation and certification of computer systems. This schema was developed based on NIST guidance developing accreditation and certification packages. This schema is a typical example of requirements for development of document types in XML.

SDLC Hierarchy
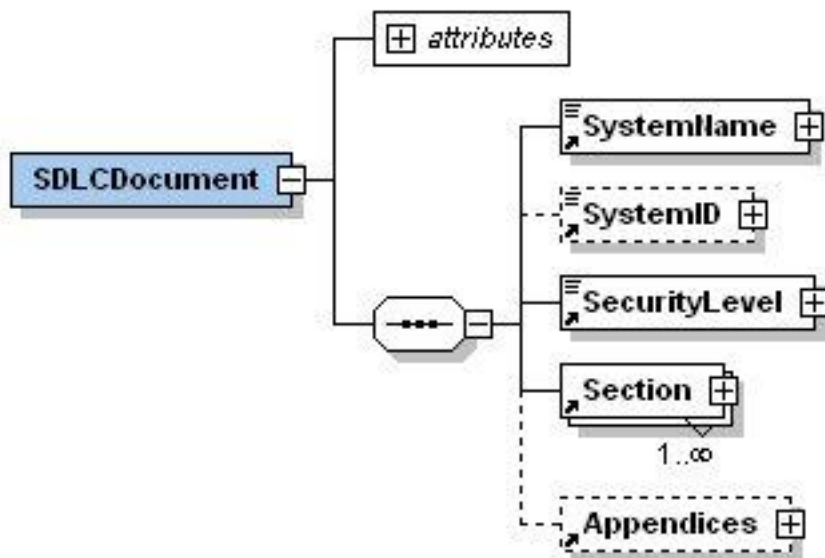
Slide 14: ATD1

User-defined attributes SHOULD NOT be used.

- Attributes are often used at the element level for metadata.
- ATD1 will eliminate the ability to add metadata at the element level
- Content management systems, such as Documentum, use attributes as a mechanism for access to documents and/or chunks of documents. Validation requires these attributes to be included in the schema.

SDLC Top Level Attributes

Slide 15: Attributes Used Effectivity

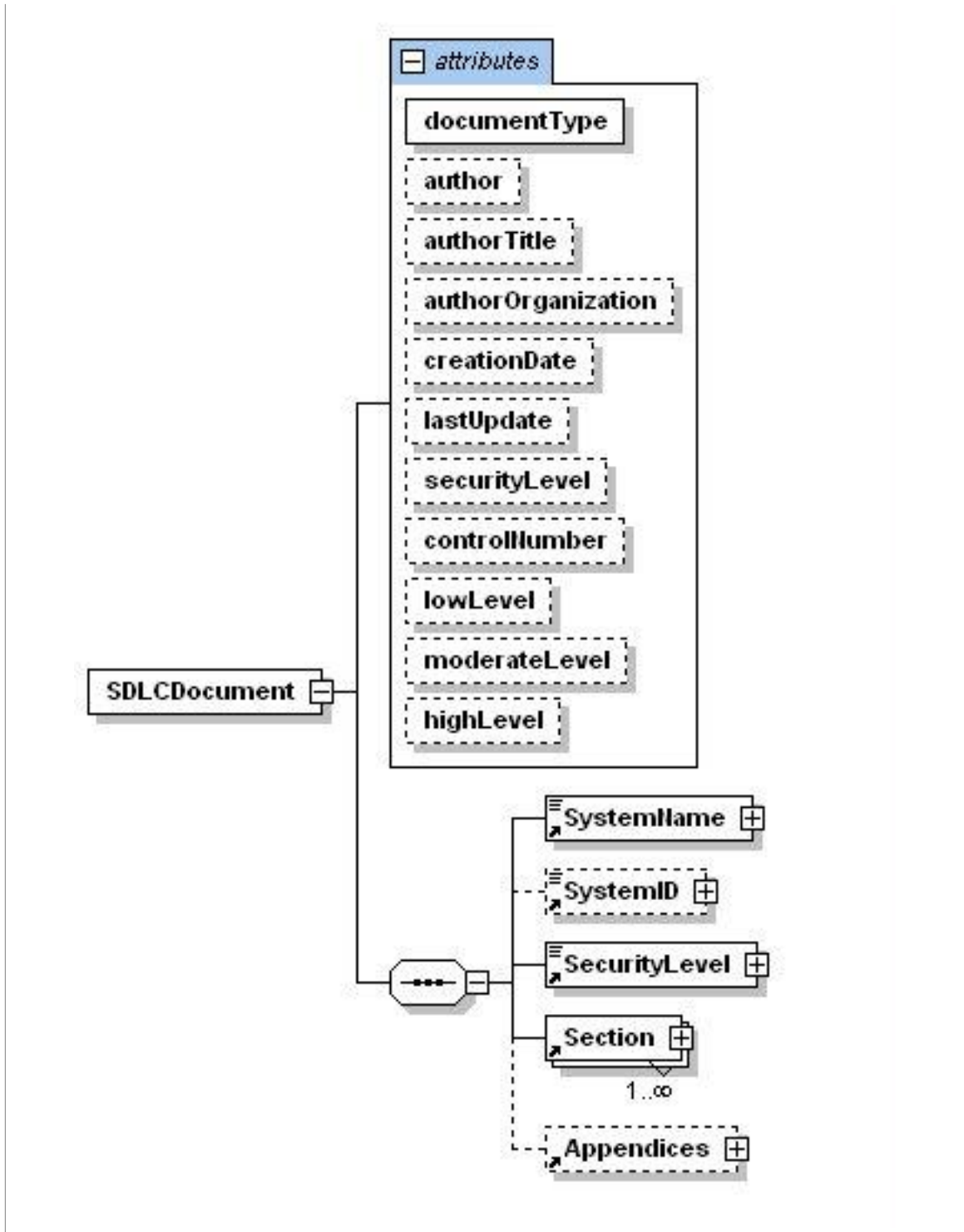In manufacturing and technical documentation **effectivity** is very important. Effectivity means that the information will changed based on type of equipment, components etc.

Effectivity is used for any piece of equipment. For example, two F18 may use different manufacturing parts. Maintenance instructions will change based on the specific part in a specific airplane. This information is extremely important to capture.

ATA (Airline Transportation Association) 2100 captures effectivity as attributes.

Example of effectivity

```
<procedure>
<step effectivity="ISDN">You can connect to an analog
   telephone port (an R-interface port) on an ISDN terminal adapter. For more
   help and to request an R-interface port, contact the ISDN provider.
</step>
<step effectivity="DSL">If you use a DSL (Digital Subscriber Line) phone service,
   you can connect to a DSL filter or router that will allow for some analog
   (rather than digital) use.
</step>
</procedure>
```

Slide 16: ELD6

Empty elements SHALL NOT be declared except for reference elements and Xlink elements, which MUST be approved by the cognizant FNC and BSC.

- EMPTY elements are traditionally used for graphics, multimedia, etc.
- The bureaucracy of getting FNC and BSC approval will greatly increase development time.
- People will circumvent this rule by creating a 'string' type and not plan on entering the data which will cause confusion on the user of the schema.
- Toolsets sometimes dictate how elements are created.

Empty Elements - CommentStart/CommentEnd

Slide 17: GNR1

DON XML element, attribute, and type names MUST be in the English language, using the Oxford English Dictionary for Writers and Editors (Latest Ed.). Where both American and English spellings of the same word are provided, the American spelling MUST be used

Slide 18: GRS4 and GRS5

Abbreviations and acronyms MUST be submitted to an FNC for approval.

The abbreviations and acronyms list approved by the BSC and FNC MUST be used.

- Acronyms are domain-specific
  - SME - Subject Matter Expert/Small and Medium Enterprise
  - CALS - Computer-Aided Logistics/Continuous Acquisition and Lifecycle Support/Commerce At Lightspeed/Computer-Aided Acquisition and Logistic Support /College of Agriculture and Life Sciences
  - BSR - BIOpolymer Markup Language/Business Semantic Repository/Business Service Request
  - DISA - Data Interchange Standards Association/Defense Information Systems Agency
  - OASIS - Open Applications Group Interchange Specification/Organization for the Advancement of Structured Information
  - RSS - RDF Site Summary/Really Simple Syndication.
  - RTF - Rich Text Format/Result Tree Fragment
  - UBL - Unified Business Language/Usama Bin Laden

Slide 19: GRS6

DON XML element, attribute, and type names MUST be in singular form unless the concept itself is plural (example: goods).

- This relates to grouping of like pieces of information. If users can't use the (s) to group then naming guidance should be provided for grouping of like information.

<article-categories> - Source NIH Journal Schema
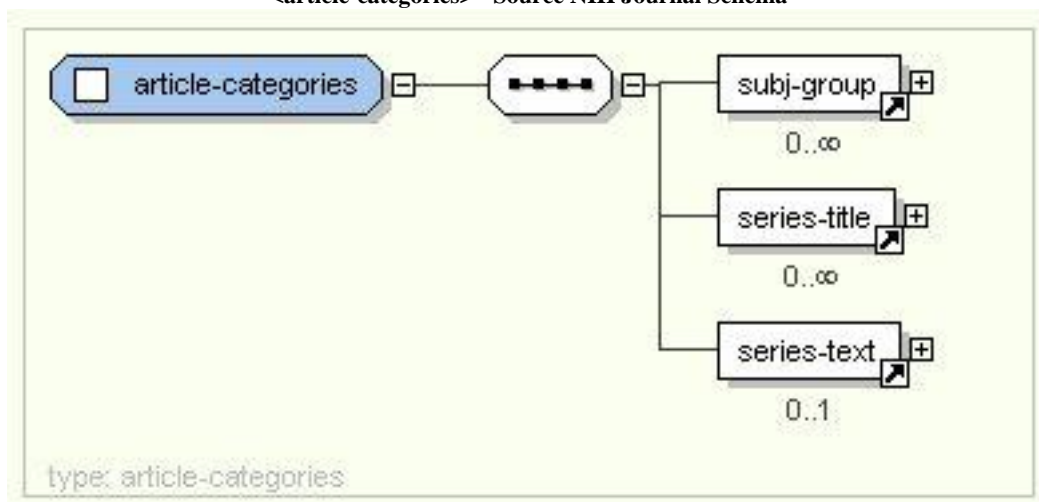
Reference Documents
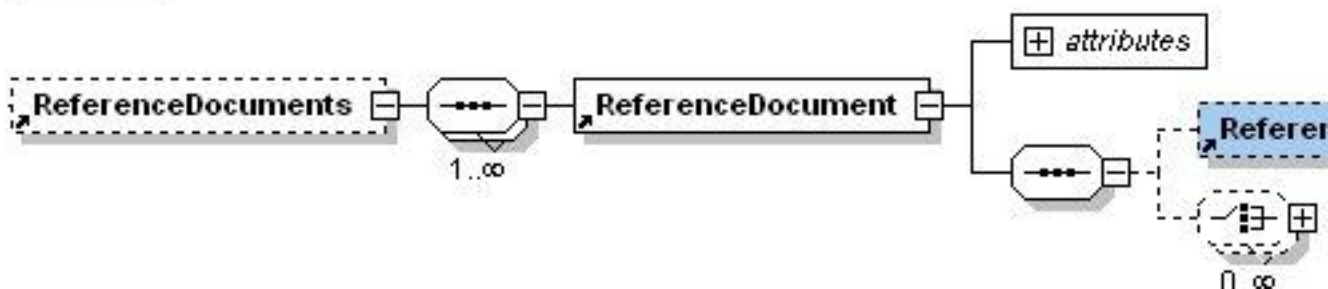
Slide 20: GRS7 and GRS8

The UpperCamelCase (UCC) convention MUST be used for naming elements and types.

The lowerCamelCase (LCC) convention MUST be used for naming attributes

- Legacy and industry standards DTDs/Schema's do not adhere to this rule
    - CALS
    - AECMA 2100
    - Legislation (Bills)
    - Docbook
    - MathML
    - XHTML

No guidance in NDR for integrating legacy standards and standard industry components, i.e., Mathml

Slide 21: MDC4

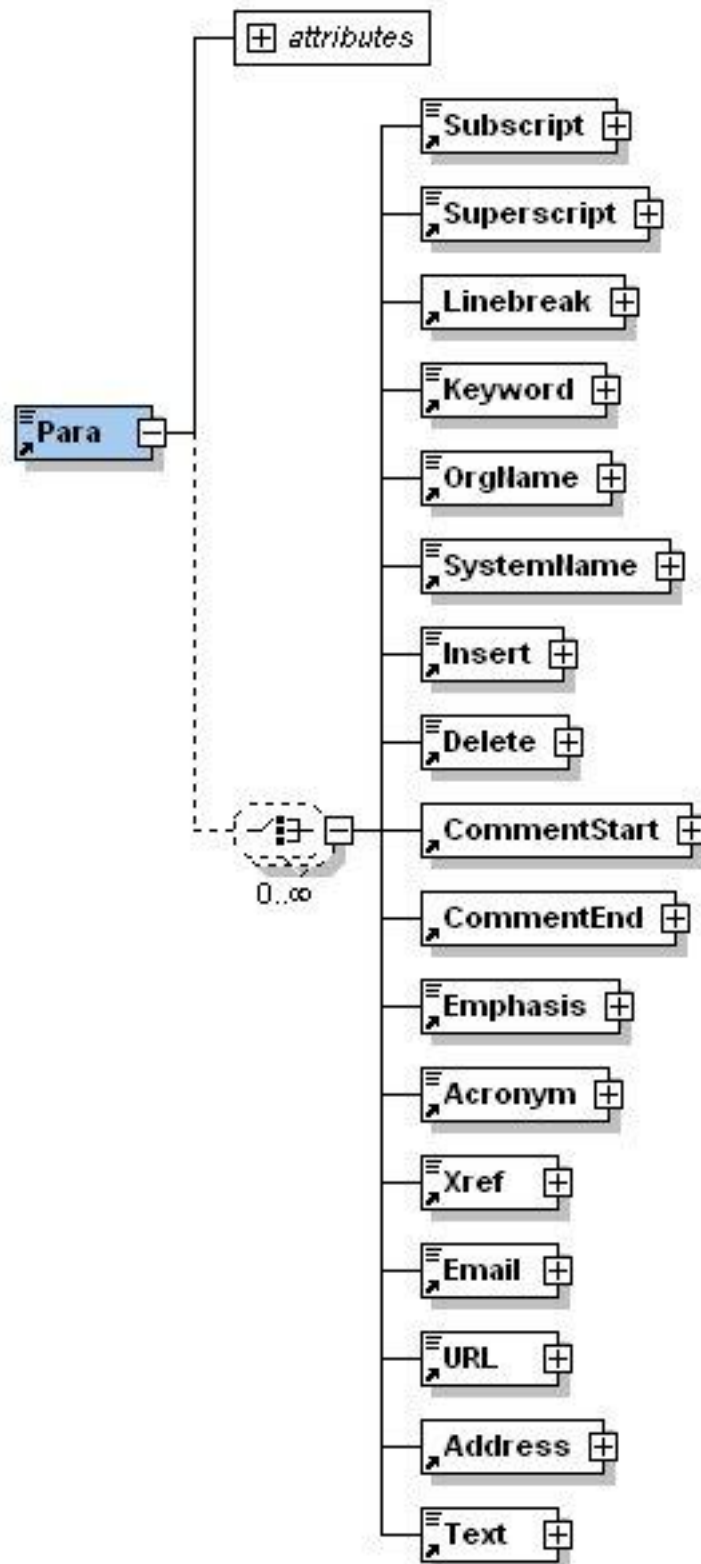Mixed content MAY only be used when an XML schema component is defined by a namespace from a BSC-approved business standard (e.g. XHTML).

The example below is typical of paragraph data with mixed content.

Mixed Content Model

Slide 22: Mixed Content Example

OMB Interior Final Output

For expenses necessary for management, protection, and development of resources and for construction, operation, and maintenance of access roads, reforestation, and other improvements on the revested Oregon and California Railroad grant lands, on other Federal lands in the Oregon and California land-grant counties of Oregon, and on adjacent rights-of-way; and acquisition of lands or interests therein including existing connecting roads on or adjacent to such grant lands; [$105,165,000] *$110,709,000*, to remain available until expended: *Provided*, That 25 percent of the aggregate of all receipts during the current fiscal year from the revested Oregon and California Railroad grant lands is hereby made a charge against the Oregon and California land-grant fund and shall be transferred to the General Fund in the Treasury in accordance with the second paragraph of subsection (b) of title II of the Act of August 28, 1937 (50 Stat. 876). *(Department of the Interior and Related Agencies Appropriations Act, 2002; additional authorizing legislation required.)*

Tagged Example of Appropriations Paragraph

```
<appropriations-para>
   <para>For expenses necessary for management,
         protection, and development of resources
   and for construction, operation, and maintenance
   of access roads, reforestation, and other
   improvements on the revested Oregon and California
   Railroad grant lands, on other Federal lands in the
   Oregon and California land-grant counties of Oregon,
   and on adjacent rights-of-way; and acquisition of
   lands or interests therein including existing connecting roads
   on or adjacent to such grant lands;
         <deleted-phrase>$105,165,000</deleted-phrase>
         <added-phrase>$110,709,000</added-phrase>,
         to remain available until expended: Provided, That 25
   percent of the aggregate of all receipts during the current
   fiscal year from the revested Oregon and California Railroad
   grant lands is hereby made a charge against the Oregon and
   California land-grant fund and shall be transferred to the
   General Fund in the Treasury in accordance with the second
   paragraph of subsection (b) of title II of the Act of August
   28, 1937
   (<cross-reference
      href="http://www.access.gpo.gov/uscode/title50a/50a_4_1_.html">
      50 Stat. 876</cross-reference>).
```

```
            </para>
            <citation>(Department of the Interior and Related
            Agencies Appropriations Act, 2002; additional authorizing
            legislation required.)
            </citation>
      </appropriations-para>
```

Slide 23: Mixed Content In A Non-Mixed Model

Example 1 - actual solution

```
<footnote>
<para>&it;Compare&eit; Limerick Ecology Action, Inc. v. Nuclear Regulatory
    Comm'n, 869 F.2d 719, 743 (3d Cir. 1989) (Nuclear Regulatory Commission, an
    independent regulatory agency, is not bound by CEQ regulations it had not
    adopted), &it;with&eit; Steamboaters v. Federal Energy Regulatory Comm'n,
    759 F.2d 1382, 1392&ndash;93 &amp; n.4 (9th Cir. 1985) (Federal Energy Regulatory
    Commission, an independent regulatory agency, is bound by CEQ regulations).
</para>
</footnote>
```

Another Solution

```
<footnote>
<para><keyword>Compare</keyword>
    <text>Limerick Ecology Action, Inc. v. Nuclear Regulatory
        Comm'n, 869 F.2d 719, 743 (3d Cir. 1989) (Nuclear Regulatory Commission, an
        independent regulatory agency, is not bound by CEQ regulations it had not
        adopted),
    </text>
    <keyword>with</keyword>
    <text>Steamboaters v. Federal Energy Regulatory Comm'n,
        759 F.2d 1382, 1392&ndash;93 &amp; n.4 (9th Cir. 1985) (Federal Energy
        Regulatory Commission, an independent regulatory agency, is bound by CEQ
        regulations).
     </text>
</para>
</footnote>
```
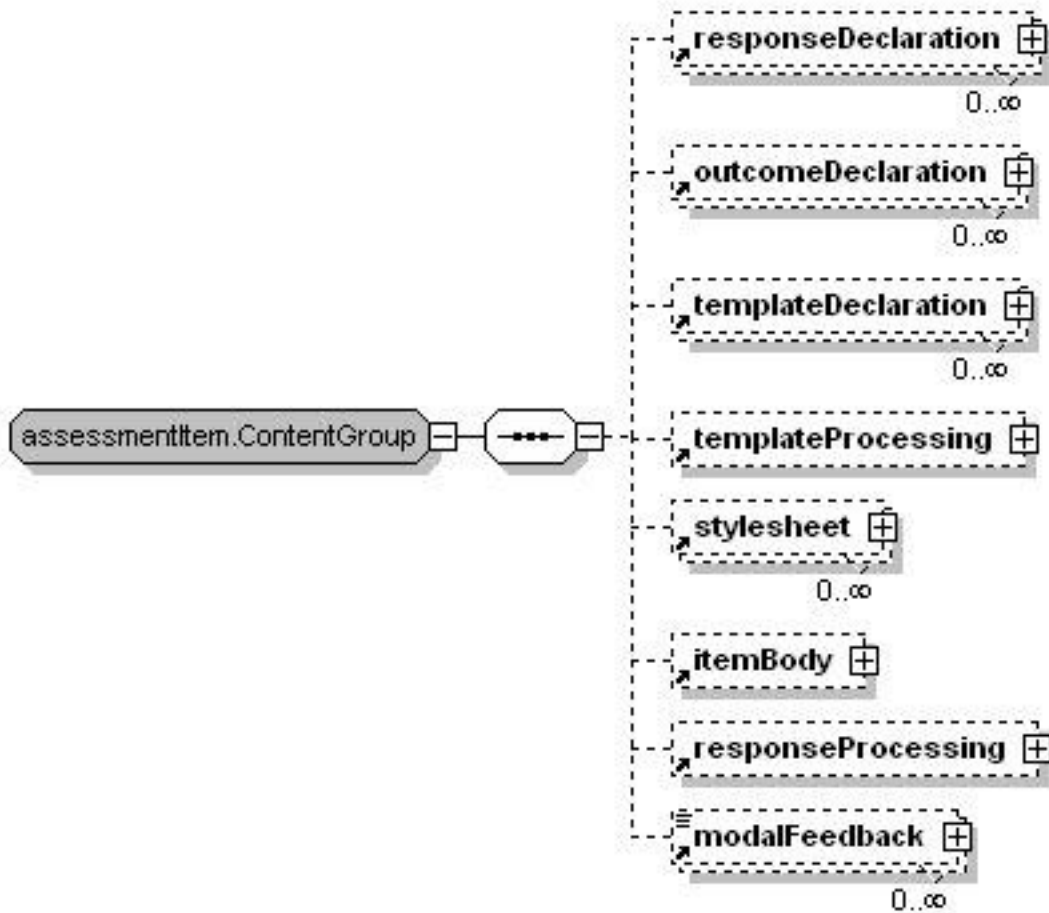
Slide 24: Need Guidance for Technical Documents

- NDR provides no guidance for XML converted legacy documents
    - AMSEC LLC MPP - AMSEC LLC Multi-Purpose Publishing
    - EOSS - Engineering Operational Sequencing System
    - MIL-M-38784 - Manuals, Technical: General Style and Format Requirements
    - MIL-M-81310 - Manuals, Technical: Airborne Weapons/Stores Loading/Weapons Assembly/Support Equipment Configuration
    - MIL-PRF-87269 - Interactive Electronic Technical Manuals DataBase (IETMDB)
    - MIL-STD-3001 - NAVAIR Work Package Technical Manual
    - NAVSEA Class 2 - NAVSEA Class 2 Electronic Technical Manual (Updated May 2005)
    - NAVSEA OD IEP - NAVSEA Ordnance Document (OD) Interactive Electronic Procedure (IEP) initiative
    - Navy ETM XML - Navy Electronic Technical Manual XML (Updated May 2005)
    - PMS - Planned Maintenance System
- No guidance for use of industry standard components
    - MathML
    - Metadata Standards (Dublin Core/Marc/ONIX, etc)
    - Table Models (OASIS-Open nee CALS)
    - AECMA S1000D

Air Force Guidance: *In March 2003, DFSG/SB (formerly MSG/MM) issued interim guidance for USAF Interactive Electronic Technical Manuals (IETMs). The guidance indicated that future USAF IETMs will be based on S1000D, The international specification for technical data utilizing a common source database. At that time, S1000D was already a fully capable linear technical data specification that was undergoing dramatic enhancements to additionally satisfy non-linear, also called IETM, requirements.*

IMS Question and Test

Slide 25: Questions