Exchanging eContracts Using XHTML 2.0 and Modular XHTML

OASIS/LegalXML/eContracts

eContract Structure Sub-committee

Working Draft, Version 1

Sub-committee members:

Jason Harrop, jharrop@speedlegal.com
Dr Laurence Leff, D-Leff@wiu.edu
John McClure, jmcclure@hypergrove.com
Mary McRae, mmcrae@dmsi-world.com
Peter Meyer, pmeyer@elkera.com.au

Background

The Structural Subcommittee was organized in February 2004 to examine XHTML 2.0 as an implementing technology for the representation of the document structure common to many legal contracts. Initial discussions focused on the fundamental architecture of legal clauses, resulting in the unanimous recommendation that the eContracts TC adopt the "grammatical paragraph" model, where all text content of any "clause" is to be located within <p> elements only.

Our subsequent work focused on additions to, changes of, and deletions from, the elements defined by the XHTML 2.0 language. Our functional requirements were established by the TC relating to first, the structure of an XHTML file representing a contract and second, to more general objectives of interested stakeholders. Specifically, the Subcommittee did not examine the use of XHTML 2.0 to meet TC requirements for "semantic" markup, that is, XML elements or attributes identifying the functional purpose of text within a digital contract, e.g., the contract's title, parties, dates, terms, or any metadata about the production, acceptance, or management of the contract. The scope was delineated as the "structure" of the contract.

Note: Our discussions to date have been based on the "penultimate" XHTML 2.0 working draft of July 2004.

Table of Contents

Document Structure

Basic Structure. Below is the markup for the second section of this document. The section contains a titled group of paragraph(s) within a section, and two titled groups of subsections. The document's organization is typical to formalized contracts.

```
<instrument>
  <section property="OperativeClause"/>
  <section property="OperativeClause">
    <nr>2.</nr>
    <h>Captions</h>

    <h name="group1">Background</h>
    <p name="group1">Captions are used to identify specific content within...</p>
    <p name="group1">The XHTML 2.0 language introduces the <h>...</p>
    <p name="group1"> The numeric part of a caption may itself consist of...</p>

    <h name="group2">Issues Examined</h>
    <section name="group2"">
      <nr>2.1</nr>
      <p>No element exists in XHTML 2.0 to identify the numeric part ...</p>
    </section>
    <section name="group2"">
      <nr>2.2</nr>
      <p>Auto-numbering and limited text generation is a feature of ...</p>
    </section>

    <h name="group3">Recommendations</h>
    <section name="group3"">
      <nr>2.3</nr>
      <p>Designate a new inline element "
        <nr> - the container for...</p>
    </section>
    <section name="group3"">
      <nr>2.4</nr>
      <p> Require numeric captioning for each
        <section> element...</p>
        </section>
    </section>
  </section>
</instrument>
```

1. **Clauses**

Background

The term "clause" represents any numerically captioned block of text within the operative text of a legal instrument. No consensus exists in the international legal community about the name for this text block.

The XHTML 2.0 language introduces the <section> element as a new block-level element, with no designation about its functional use. The <section> element is a recursive element, meaning that a <section> may contain <section> elements. The other block-level, flowing-text recursive element provided by XHTML is the <div> element.

Issues Examined

1.1   The <section> content model breaks the "grammatical paragraph" model requirement by allowing #PCDATA, <div>, <table>, and list elements as components.

1.2   Groups of clauses, indicated by a preceding heading, must be represented in a contract structure. The XHTML 2.0 mechanism for grouping elements, the @name attribute, was evaluated for application to groups of clauses.

Recommendations

1.3   Establish the <section> element as the markup element for a numbered clause.

1.4   Allow multiple <section> elements within an <instrument> element. Require the use of an "OperativeClause" and "InoperativeClause" keywords to identify the <section> element that contains an operative or inoperative clause for the instrument.

1.5   Require that all clauses be numerically captioned, and optionally titled, to (a) functionally distinguish clauses from uncaptioned grammatical paragraph content (b) enable the creation of a table of contents and (c) generate standardized reference strings to one or more clauses.

1.6   Remove #PCDATA, <table>, <div>, and list elements from the <section> element content model.

1.7   Identify hypertext references to a clause using the "ClauseReference" keyword; to a group of clauses using the "ClauseGroupReference" both to provide for standardized reference strings and a rich User Experience.

1.8   Identify the membership of a clause in a specific clause group by using the @name attribute on the <section> element. The value of this attribute is assigned by an author, which should correlate to the value of a @name attribute on an associated <h> element located within the <section> element's parent <section> element.

2. **Captions**

Background

Captions are used to identify specific content within a legal instrument, each consisting of a number and a title, depending on the item captioned. All clauses have, at a minimum, a numeric caption. Tables and images are optionally captioned with a number and or a title. Each item either appended or prepended is captioned, at a minimum, with a title.

The numeric part of a caption may itself consist of a text or image prefix to a structured number reflecting at least the sequence of its captioned material, and optionally its depth within the clause tree. The numeric part may optionally be suffixed by a "period". For certain contexts, additional suffixing text strings or images are likely required.

The XHTML 2.0 language introduces the <h> element as a replacement for the <h1> to <h6> elements.

Issues Examined

2.1   No element exists in XHTML 2.0 to identify the numeric part of a caption. Numeric captions are an established requirement for clauses, tables, images, and appended and prepended material.

2.2   Auto-numbering and limited text generation is a feature of CSS-2, a styling standard inconsistently implemented across tools. Our ability to meet our print and display fidelity requirements must not be predicated on complete implementation of CSS-2 by User Agents.

2.3   Captions for grouped paragraphs and grouped clauses are a mandatory requirement. XHTML provides a grouping mechanism, the @name attribute.

Recommendations

2.4   Designate a new inline element " <nr> " the container for numeric caption text strings.

2.5   Require numeric captioning for each <section> element

2.6   Disallow styling-based generation of the content of either a numeric or a title caption.

2.7   Allow <section> elements to contain a single <nr> element and multiple <h> elements.

3.   **Tables**

Issues Examined

3.1   The <table> element is allowed within various XHTML 2.0 elements that are outside the scope of a <p> element, breaking the "grammatical paragraph" model requirement.

3.2   The <table> element does not support captioning. Legal instruments contain captioned tables, refer to captioned tables, and index captioned tables.

3.3   The <table> element supports styling attributes that jeopardize the fidelity of a contract's representation. Styling is already provided through mechanisms external to the <table> element and its child elements.

3.4   The CALS/OASIS table markup model is not supported by XHTML 2.0; this is an important standard from our organization.

3.5   Table bodies, rable rows, and table cells cannot be captioned by XHTML 2.0 elements. Legal instruments reference material located in a table by row-group, by row, and by cell. Without captioning, standardized reference strings cannot be generated.

Recommendations

3.6   Remove the <table> element from the content models for <body>, <section>, and <div> elements, to meet the "grammatical paragraph" requirement.

3.7   Add the <nr> and <h> elements as optional components of the <table> element's content model, to support table captioning.

3.8   Remove the @rules attribute from the <table> element, to promote presentation fidelity.

3.9   Use standard XML Namespace conventions, to support inclusion of CALS/OASIS table markup in an XHTML file. Identify or create XSLT stylesheets to transform a CALS table to XHTML 2.0 markup.

3.10   Add the <nr> and <h> elements as optional components of the <tbody>, <tr>, and <td> content models, to provide captioning for table row groups, table rows, and table cells.

3.11   Identify hypertext references to a table using the "TableReference" keyword; to a row-group using the "RowGroupReference" keyword; to a row using the "RowReference" keyword; and to a cell using the "CellReference"" keyword, to provide for standardized reference strings and a rich User Experience.

4.   **Objects**

Background

The WD says: "The Object Module provides elements for general-purpose object inclusion; this includes images and other media, as well as executable content. .. When this module is used, it adds the object element to the Inline content set of the Inline Text module."

The <img> object from XHTML 1 is functionally replaced by the <object> element.

Issues Examined

4.1   The <object> element is allowed within various XHTML 2.0 elements that are outside the scope of a <p> element, breaking the "grammatical paragraph" model requirement.

4.2   The <object> element does not support captioning. Legal instruments contain captioned images, refer to captioned images, and index captioned images.

4.3   The <object> element allows executable content, i.e., a reference to an executable program. Processing by this program may jeopardize the fidelity of a contract presented some time following its acceptance by parties.

Recommendations

4.4   Remove the <object> element from the content models for <body>, <section>, and <div> elements, to meet the "grammatical paragraph" requirement.

4.5   Add the <nr> and <h> elements as optional components of the <object> element's content model, to support image captioning.

4.6   Disallow executable programs anywhere within the boundaries of the <instrument> element, to ensure the integrity of the contract following its acceptance by parties.

4.7   Identify hypertext references to an image using the "FigureReference" keyword, to provide for standardized reference strings and a rich User Experience.

5.   **Appended Material**

Background

Section 8.4 of the WD says: "The div element, in conjunction with the id and class attributes, offers a generic mechanism for adding extra structure to documents. This element defines no presentational idioms on the content. Thus, authors may use this element in conjunction with stylesheets , the xml:lang attribute, etc., to tailor XHTML to their own needs and tastes."

Issues Examined

5.1   No direct support for material appended or prepended to the primary content of a legal instrument exists in XHTML 2.0. Numerous legal instruments require multiple front- matter sections including, for example, cover pages, contents table pages, and summary pages; and multiple back-matter sections including, for example, subordinate legal instruments, generic documents, and index pages.

5.2   Multiple documents and structured documents must be able to be represented within, at least, the material appended to the main portion of a legal instrument.

5.3   The XHTML 2.0 <div> content model does not require captioning of any kind. The <section> element requires numeric captioning; the <p> element allows no captioning; and all prepended and appended material requires caption titles. These titles are used for generation of reference strings relating to the material.

Recommendations

5.4   Use the <div> element as the container for any material appended or prepended to the primary content of the legal instrument. The <div> element is recursive, meeting the need for embedding one or more documents (parts) within appended or prepended material.

5.5   Add the <nr> and <h> elements to the <div> element content model, to support captions and table of contents generation.

5.6   Remove the <div> element from the content model of the <body> and <section> elements; retain the <div> element within a table cell element content model.

5.7   Identify hypertext references to prepended material using the "FrontMatterReference"" keyword, to appended generic material using the "AttachmentReference" keyword; and to appended subordinate legal instruments using the "SubInstrumentReference" key-word, to facilitate the generation of standardized reference strings and to provide a rich User Experience. The "InstrumentReference" keyword is used for hypertext references to documents external but related to the instrument, that is, not part of the instrument element.

5.8   Add the <instrument> element as an optional component of the <div> content model, to support appended subordinate legal instruments.