

# Oasis Legal XML eContracts TC

**Document Title:** Evaluation of OpenDocument 1.0 (2005-05-01)  
(<http://www.oasis-open.org/office/v1.0>)  
**Author:** Rolly Chambers  
**Version:** Draft 0.01  
**Date submitted:** 2005-06-10

## 1. Summary Comparison of OpenDocument XML Schema against Schema Evaluation Criteria Draft 0.01 (2005-05-22)

The OpenDocument 1.0 Standard became an approved Oasis standard as of May 1, 2005. The OpenDocument 1.0 standard defines a RelaxNG XML schema for office documents, including text documents. It is based on OpenOffice.org XML, and will be used in OpenOffice, a suite of word processing, spreadsheet, and presentation applications substantially similar to Microsoft Office.

Jason Harrop, a former member of this TC previously evaluated the OpenOffice DTD, the predecessor of the OpenDocument 1.0 Relax NG schema, and reported the results in a paper entitled “The Clause Model Pattern – A Survey of Existing DTDs” (2003-07-18). The conclusion was that the OpenOffice DTD did not provide for grammatical paragraphs which can contain lists. In the OpenOffice DTD, lists are parent and sibling elements of paragraphs but cannot occur as children of paragraphs. This structure is carried forward in OpenDocument 1.0.

That the OpenDocument XML schema does not allow lists to occur within paragraphs means that it does not satisfy this TC's requirements for grammatical paragraph structures, which must contain lists. Consequently, the OpenDocument XML schema does not meet this TC's requirements for object representation of paragraphs.

The numbered rows in the table below correlate to the criteria in the Schema Evaluation Criteria Draft 0.01 (2005-05-22) by Peter Meyer of this TC.

<i>“Clause” Pattern</i>	<i>Candidate Elements</i>	<i>Comments</i>
2. Generic numbered object container/representation	<text:section>	- contains paragraph level content (text:h and text:p), but does not enforce sequence  -<text:section> must have a text:name attribute that uniquely identifies the <text:section>

<i>“Clause” Pattern</i>	<i>Candidate Elements</i>	<i>Comments</i>
Heading?	<text:h>*	<ul style="list-style-type: none"> <li>- &lt;head&gt; is optional for &lt;text:section&gt;</li> <li>- #PCDATA allowed (mixed content) in &lt;text:h&gt;</li> <li>- &lt;text:h&gt; is a sibling of &lt;text:p&gt;</li> <li>- &lt;text:h&gt; can precede or follow &lt;text:p&gt; (sequence is not enforced)</li> </ul>
Paragraph+	<text:p>*	<ul style="list-style-type: none"> <li>- &lt;text:p&gt; cannot contain lists (&lt;text:p&gt; would need to be re-defined)</li> </ul>
Hierarchy	<office:document> <office:body> <text:section> <text:h>* or <text:p>* <text:span>* <text:keywords>*	<ul style="list-style-type: none"> <li>- no phrase element;</li> <li>&lt;text:span&gt; is a candidate</li> </ul>
Additional infrastructure within <text:section> and <text:p>	<ul style="list-style-type: none"> <li>- &lt;text:section&gt; can contain tables</li> <li>- &lt;text:p&gt; can contain &lt;draw:image&gt;</li> <li>- version tracking</li> </ul>	

<i>Consideration</i>	<i>For</i>	<i>Against</i>
3. Metadata	<office:meta> uses several Dublin Core metadata elements	<ul style="list-style-type: none"> <li>- &lt;office:meta&gt; contains metadata in a separate file from &lt;office:document&gt; (in OpenOffice, the separate files are “zipped” together into a package)</li> <li>- no clearly defined mechanism for extending &lt;office:meta&gt;</li> </ul>

<i>Consideration</i>	<i>For</i>	<i>Against</i>
4. Processing technologies/license terms	<ul style="list-style-type: none"> <li>- open source</li> <li>- allows use of XSLT stylesheets to import and export content to other XML vocabularies</li> </ul>	
5. Numbering	<text:number> can occur within <text:numbered-paragraph> and must precede <text:h> or <text:p>	- <text:number> does not occur within <text:section> (no author defined numbering of sections)
6. Complete contract documentation	Authors can explicitly represent narrative contract terms	<ul style="list-style-type: none"> <li>- no clearly defined mechanism for extending OpenDocument 1.0 XML schema</li> <li>- adding to or subsetting the OpenDocument 1.0 XML elements probably would affect the functionality of OpenOffice</li> </ul>
7. Variables	<text:variable-decl> and similar elements can contain defined variables	
8. Ease of use	- OpenDocument 1.0 and OpenOffice documentation is clear, well-developed, and up-to-date	<ul style="list-style-type: none"> <li>- extensive sets of elements for spreadsheets, graphics, and styling</li> <li>- OpenOffice is an office-suite application functioning as a word processor, spreadsheet, graphics, and presentations application</li> </ul>
9. Schema syntax		- Relax NG schema only; no DTD or W3C XML schema
10. Adaptability to contracts		- no named elements for recitals, attachments, exhibits, schedules
11. Developer/vendor support and skill set	<ul style="list-style-type: none"> <li>- Well documented Relax NG schema</li> <li>- Oasis standard</li> <li>- XSLT stylesheets for importing to and exporting from OpenOffice</li> </ul>	

## **2. Conclusions**

- OpenDocument 1.0 does not meet the requirements of this TC for grammatical paragraphs – paragraphs cannot contain lists
- OpenDocument 1.0 is used in OpenOffice, an application substantially similar to office applications like Microsoft Office and would be easy for document authors to use to create XML content
- OpenOffice combines XML content, stylesheets, metadata, etc. into a ZIP file called a package. The XML content must be extracted with a ZIP application for further processing
- There is no clearly defined mechanism for extending or creating a subset of the OpenDocument XML schema
- The OpenOffice 1.0 XML vocabulary is extensive and developing a standard XSLT stylesheet for importing or exporting XML content to a simpler eContracts XML vocabulary would be useful