

XHTML Template for Conference Submissions

F R A Hopgood

Oxford Brookes University

bhopgood@brookes.ac.uk

ABSTRACT

This paper defines the IW3C2 Conference Paper Template.

Keywords

IW3C2; XML; XHTML; Template

1. Introduction

IW3C2 is committed to producing an online set of Proceedings for the complete set of WWW Conferences with the correct interlinking between volumes. IW3C2 also has a strong commitment to Web Standards and for that reason there is a requirement that papers submitted to the Conference are in the current standard HTML format (which is currently XHTML). To aid in the production of the online version, the printed Conference Proceedings and the CD-ROM, it is important that authors produce their papers in a standard format which is presentation independent and the one chosen is *Strict* XHTML. By doing this the production process can be accelerated and this allows the date for final papers to be left as late as possible.

With the current state of word processing tools, choosing any format will create problems for some authors. We believe that to achieve the goals of the Conference Series, using Strict XHTML is the one that creates least work for a knowledgeable Web community. Producing web papers in electronic paper formats such as PDF may help the production of the printed version assuming the correct template is used, but it creates major problems for the other formats required. It also makes it difficult to reuse the material.

This paper is aimed at helping authors achieve the production of their paper in Strict XHTML by giving information concerning the tools available to help in the process and to point out the areas where problems may occur.

The overall format of the template is described followed by a description of XHTML, CSS and the tools available. The problem with mathematics is discussed in detail and some comments are made on images and tables.

2. Paper Format

Following is an example showing what a submitted paper should look like:

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "DTD/xhtml1-strict.dtd">
<html>
<head>
  <title>Insert Title Here: IW3C2 Conference Paper</title>
  <link rel="stylesheet" href="iw3c2.css" />
</head>
<body>
  <div class="meta">
    <h1 class="title">Title of Conference Paper Here</h1>
    <div class="authors">
      <div class="author">
        <h2 class="author"> Albert Other</h2>
        <h3 class="affiliation">My Institute, France</h3>
        <h2 class="email">albert@xxx.org</h2>
      </div>
      <div class="author">
        <h2 class="author"> Fred Another</h2>
        <h3 class="affiliation">His Institute, USA</h3>
        <h2 class="email">fred@yyy.org</h2>
      </div>
    </div>
    <div class="abstract">
      <h1 class="abstract">ABSTRACT</h1>
      <p class="abstract">This paper attempts to provide a template for you to
        write papers.</p>
    </div>
    <div class="keywords">
      <h2 class="keywords">Keywords</h2>
      <p class="keywords">IW3C2; XML; XHTML; XSLT</p>
    </div>
  </div>
```

```
<div class="paperbody">
<h1>1. INTRODUCTION</h1>
<p>Following is the content of the paper. Note that papers should be formatted
using standard XHTML elements (h1-h6, ul, ol, li, p etc). Do not use
presentation elements like font. The www2003.css Cascading Style Sheet will take
care of formatting the text for you. Format section headings in your paper
as in this paper, div sections for the metadata at the top. The ABSTRACT is headed by
an h1 element of class abstract and contains a p element of
class abstract. The Keywords section is headed by an h2 element of
class keywords and contains a p element of class keywords. </p>
<h2>1.1 Subheadings</h2>
<p>Subheadings are formatted as h2 elements.</p>
<h1>2. IMAGES</h1>
<p>These should be placed in the same directory as the paper and linked to directly.
Preferred format for computer graphics images is png and for real world images
is jpg. Vector graphics should use svg. Use the object element to provide
an alternative image format.</p>
<h1>3. HOW TO DO REFERENCES</h1>
<p>Local references should appear as below <a href="#r1">[1]</a>.
Hypertext references should use the
<a class="href" href="http://www.iw3c2.org/">IW3C2</a> formalism
(define the a element as of class href) shown in this example.
This is so that an XSLT transformation will allow the printed
proceedings to contain the full URLs for use by readers.</p>
</div>
<div class="references">
<h1>REFERENCES</h1>
<p class="ref" id="r1">[1]
  <span class="authors">A. Bloggs</span>,
  <span class="title">Web Navigation: Designing the User Experience</span>,
  <span class="journal">Web Journal</span>,
  <span class="details">13 (2) 1988</span>
</p>
<p class="ref" id="r2">[2]
  <span class="authors">D.E. Alpha, J.R. Beta, L.M. Gamma</span>,
  <span class="title">Web Searching: A Good Experience</span>,
  <span class="book">Web Searching</span>
  <span class="details">Academic 1988</span>
</p>
</div>
</body>
</html>
```

This would have specific styling for the paper and CD versions but they might look like this:

Title of Conference Paper Here

Albert Other
My Institute, France
albert@xxx.org

Fred Another
His Institute, USA
fred@yyy.org

ABSTRACT
This paper attempts to provide a template for you to write papers.

Keywords
IW3C2; XML; XHTML; XSLT

1. INTRODUCTION
Following is the content of the paper. Note that papers should be formatted using standard XHTML elements (h1-h6, ul, ol, li, p etc). Do not use presentation elements like font. The www2003.css Cascading Style Sheet will take care of formatting the text for you. Format section headings in your paper as in this paper, div sections for the metadata at the top. The ABSTRACT is headed by an h1 element of class abstract and contains a p element of class abstract. The Keywords section is headed by an h2 element of class keywords and contains a p element of class keywords. Appendix A lists a set of keywords used at previous conferences.

1.1 Subheadings
Subheadings are formatted as h2 elements.

2. IMAGES
These should be placed in the same directory as the paper and linked to directly. Preferred format for computer graphics images is png and for real world images is jpg. Vector graphics should use svg. Use the object element to provide an alternative image format.

3. HOW TO DO REFERENCES
Local references should appear as below [1]. Hypertext references should use the IW3C2 formalism (define the a element as of class href) shown in this example. This is so that an XSLT transformation will allow the printed proceedings to contain the full URLs for use by readers.

REFERENCES
[1] A. Bloggs, Web Navigation: Designing the User Experience, Web Journal, 13 (2) 1988
[2] D.E. Alpha, J.R. Beta, L.M. Gamma, Web Searching: A Good Experience, Web Searching Academic 1988

Using XSL-FO, it is relatively simple to transform the XHTML into the appropriate layout required by the publisher for the printed version of the Proceedings.

The following sections describe methods to achieve the desired format shown above.

3. XHTML

The current W3C standard for HTML is *XHTML*. That is not as big a difference as you might think from what you are doing at the moment. The *X* indicates that the HTML is well-formed XML and that brings with it the ability to transform and reformat the papers using the standard XML tools. In XHTML, marked up information consists of a set of elements that have the general form:

```
<xyz>Some content</xyz>
```

Both the start and end tags *must* be lowercase, and the end tag must be present. Tags must also be correctly nested. This would be illegal:

```
<xyz>Some <abc>more content<def>added</abc></def></xyz>
```

Either the **def** element must be completely inside the **abc** element or completely outside it.

Elements may have attributes and these form part of the start tag:

```
<xyz attr1="sometext" attr2='somemoretext'>Some content</xyz>
```

Attributes consist of the *attribute name* followed by the *equal sign* and a value which is a string of text inside one of two types of quotation marks (double quotes: ASCII decimal 34; single quotation marks: ASCII decimal 39). Two types are allowed in case the attribute's value contains quotation marks. The start and end quotation marks must either both be single or both be double quotes.

A shorthand is provided for element's that contain no content. For example:

```
<xyz attr1="sometext" attr2='somemoretext'></xyz>
```

may also be written:

```
<xyz attr1="sometext" attr2='somemoretext' />
```

A strict XHTML 1.0 document would have the following structure:

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "DTD/xhtml11-strict.dtd">
<html>
<head>
. . .
</head>
<body>
. . .
</body>
</html>
```

The designation of *strict* requires that styling of the document is done via CSS and not by attributes applied to the XHTML elements.

4. CSS

To achieve device independence, styling for the various versions of the Proceedings will be achieved by applying a standard CSS stylesheet to the paper. For example, the paper would be modified for the CD-ROM version by linking in a standard CSS stylesheet:

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "DTD/xhtml11-strict.dtd">
<html>
<head>
<link href="www20xxcdrom.css" rel="stylesheet" type="text/css" title="IW3C2 Standard Styling for CD-ROM"/>
. . .
</head>
<body>
. . .
</body>
</html>
```

If the author has a real need for some very specific styling for some part of his paper, this should be achieved by either linking to a separate CSS stylesheet for the paper or including a **style** element in the head of the document.

5. Tools

Some important tools for changing HTML of various vintages into XHTML are available. These are:

- **HTML Tidy:** originally from W3C but now located at <http://tidy.sourceforge.net/> will take HTML and will either list the errors or, if it can understand it, will convert it into XHTML for you. The following command will take your HTML file and convert it to XHTML:

```
tidy -f errorfile.txt -m -asxml -clean myfile.html
```

 Note this will overwrite your existing file so make a copy before you do this! A number of HTML editors (for example, HTML-Kit) contain Tidy as part of the Editor or provide a button to access it directly. Authors should check their favourite tool to see if this option is available as it may make the task of accessing Tidy easier.
- **XHTML Validator:** located at <http://validator.w3.org/> this will check a file you present it against the various versions of XHTML. You can upload your paper to it and check it against Strict XHTML and see if it passes and if not why not. Papers that are not Strict XHTML will not be accepted for the Conference unless the author has received some specific dispensation due to some unusual requirements.
- **Word:** it is possible to generate XHTML of reasonable quality starting from a Word document. You have to realise that SAVE AS into an HTML file will not do the job. That will leave a large amount of Word information in the document to allow you to move back to a Word document. Microsoft provides a utility that you can download called **HTML Filter**. You then *Export To* the **Compact HTML** format from Word. This will remove the Word specific parts of the file. If the file is then put through Tidy you will land up with a small compact XHTML document. The download site is currently <http://office.microsoft.com/downloads/2000/Msohtml2.aspx> although a search for **Microsoft HTML Filter** will find it.
- **Latex:** a good starting point is latex2html at <http://www.tug.org/mailman/listinfo/latex2html>. Note this does not produce valid XHTML at the time of writing and the author will need to put it through Tidy after doing the conversion and may need to remove local styling attributes depending on the format of the initial Latex document.
- **Open Office:** this can also be converted to HTML and run through Tidy to produce a valid XHTML document. Some manual editing may be necessary afterwards.

Note that in almost all cases the result may need some additional editorial work to get it to pass through the XHTML Strict Validation.

6. Mathematics

The preferred approach to add mathematics to XHTML is to use MathML data islands within the XHTML document. Conversion tools exist from most existing mathematical formats to MathML. Users of IE need to download the IE6 MathPlayer 1.0 MathML Plug-In from <http://www.mathtype.com/en/products/mathplayer/>. This site also has a great deal of relevant information concerning conversion tools for MathML.

The necessary IE behaviour additions are:

```
<html xmlns:m="http://www.w3.org/1998/Math/MathML">
<head>
<object id="behave1" classid="clsid:32F66A20-7614-11D4-BD11-00104BD3F987"></object>
<?import namespace="m" implementation="#behave1" ?>
</head><body>
<h2>Expanding a binomial</h2>
<p>
<m:math id="eq1">
<m:mrow>
  <m:mrow>
    <m:msup>
      <m:mrow>
        <m:mo>( </m:mo><m:mi>a</m:mi><m:mo>+</m:mo><m:mi>b</m:mi><m:mo> )</m:mo>
      </m:mrow>
      <m:mn>2</m:mn>
    </m:msup>
  </m:mrow>
</m:mrow>
<m:mo>=</m:mo>
<m:mrow>
  <m:msup><m:mi>a</m:mi><m:mn>2</m:mn></m:msup>
  <m:mo>+</m:mo>
  <m:mn>2</m:mn>
  <m:mi>a</m:mi>
  <m:mi>b</m:mi>
  <m:mo>+</m:mo>
  <m:msup><m:mi>b</m:mi><m:mn>2</m:mn></m:msup>
</m:mrow>
</m:mrow>
</m:math>
</p>
</body>
</html>
```

The latest version of Mozilla has native support so all that is required is:

```
<h2>Expanding a binomial</h2>
<p>
  <math xmlns="http://www.w3.org/1998/Math/MathML">
    <m:mrow>
      <m:mrow>
        . . . .
      </m:mrow>
    </math>
```

The W3C Math Working Group has recently released a [Universal Math XSL stylesheet \(UMSS\)](#), developed by David Carlisle to hide the IE-specific declarations:

```
<?xml-stylesheet type="text/xsl" href="mathml.xsl"?>
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>...</head>
  <body>
    <h1>Example</h1>
    ...
    <math xmlns="http://www.w3.org/1998/Math/MathML">
      <mi>x</mi><mo>+</mo><mn>3</mn>
    </math>
  </body>
</html>
```

Submitting in this form is acceptable. Download the `mathml.xsl` and other transformation files to your local system if using IE. These files need not be sent with the paper.

Rendering the MathML as an image is also acceptable:

```
<p>

</p>
```

In the example above, the mathematical formula is displayed as a GIF image. The result would look something like:

$$\frac{x}{e^x - x} = 1 - \frac{x}{2} + \frac{B_1 x^2}{2!} - \frac{B_2 x^4}{4!} + \frac{B_3 x^6}{6!} - \dots \quad |x| < 2\pi$$

Utilities to convert MathML to images can be found at <http://www.mathmlcentral.com/Tools/MSP/IntegrateMathML>.

7. Images

If you add images in GIF, PNG or JPEG format, it is probably sensible to use the `img` element. Be sure to make it a valid XHTML element and make sure it has an `alt` attribute. For example, the preferred markup is:

```
<div class="image">
<p class="image"></p>
<p class="caption">The IW3C2 Logo</p>
</div>
```

This would be equivalent to adding Figure 1.



Figure 1. The IW3C2 Logo

For SVG graphics use:

```
<div class="image">
<object width="150" height="150" data="iw3c2.svg" type="image/svg+xml">

</object>
<p class="caption">Figure 2. The IW3C2 Logo in SVG</p>
</div>
```

The SVG will be used if the browser supports SVG and otherwise the PNG image. The SVG should be defined using a `viewBox` attribute to establish the local coordinate system. SVG implementations like Batik will do the conversion to PNG for you. An `img` alternative should always be provided. Figure 2 is an example.



Figure 2. The IW3C2 Logo in SVG

8. Tables

The `table` element must have a `summary` attribute. Try and keep tables simple. A good format for a table with headings and an initial header column would look something like:

```
<div class="table">
<p class="caption">The Table Title</p>
<table summary="My Test Table">
<col width="20%" />
<col width="30%" />
<col width="30%" />
<thead>
<tr>
<th>First</th><th>Second</th><th>Third</th>
</tr>
</thead>
<tbody>
<tr>
<th>Heading</th><td>Data</td><td>Data</td>
</tr>
<tr>
<th>Heading2</th><td>Data2</td><td>Data2</td>
</tr>
</tbody>
</table>
</div>
```

which would appear as:

First	Second	Third
Heading	Data	Data
Heading2	Data2	Data2

The default stylesheet presents `th` elements centred if they appear in the `thead` of the table and left adjusted in the `tbody` of the table. Tables of class `data` have the `td` elements right adjusted.

9. Non-English Text

For non-English text, please add the appropriate `lang` attribute. For example:

```
<p>The phrase <span lang="fr">au revoir</span> is widely used in England and the USA</p>
```

Appendix A

Keywords for Technical Areas

Below is a suggested set of keywords and their sub-classes that have been used in previous Conferences.

- User Interface
 - Document Object Model
 - Multisensory Interfaces
 - Distributed objects
 - Agent Technologies
 - Wireless and mobility
 - Collaborative systems
 - Virtual reality
 - Dynamic content
- Browsers and Tools
 - Archiving and Website Tracking
 - Browsers on Mobile Devices
 - Middleware and Browser Interactions
 - Novel Browsing Paradigms
 - Tools for Web Personalisation
 - Web Navigation Strategies
 - Web Visualization
- Generic Applications
 - MathML
 - Graphics: SVG, Web3D
 - Image, audio and video applications
 - Non-standard document types
- Culture and Society
 - Rights Management
 - Cultural Communities
 - Digital Arts
 - Digitising Collections
 - Description of Resources and Collections
 - Government Policies and Standards
 - Models for Creative and Scholarly Self-publishing
 - Multi-culturalism and Multi-linguality
 - Virtual Institutions- including economic models
- E-commerce
 - Automated negotiation and bargaining
 - Computational markets
 - Dynamic pricing
 - E-markets
 - E-commerce standards
 - Online auctions
 - Personalization services
 - Shopbots
- Security
 - Access control
 - Digital signatures
 - Privacy
 - Public key infrastructure
 - Security in content distribution networks
 - Denial of service protection
- Hypermedia
 - Link architecture, management, authoring
 - Navigation
 - Link services
- Document Standards
 - XHTML and XML
 - Markup and Data Formats
- Performance, Reliability, Scalability
 - Caching
 - Capacity planning
 - Content distribution
 - Dependability
 - Fault tolerance
 - High availability
 - Load balancing
 - Proxy and server performance
 - Quality of service
 - Scalability
 - Traffic characterization
 - Flash crowds
- Practice and Experience
 - Case studies
 - Critical appraisal of systems
 - Large scale projects
 - Novel applications
 - User behaviour
 - Web site evolution
 - Web engineering
- Searching Querying and Indexing
 - Storage infrastructure
 - Indexing
 - Searching and ranking
 - Data mining
 - Query processing and optimization
 - Query expansion
 - PageRank
 - Distributed searching
 - Similarity searching
 - Answer extraction
- Web and Education
 - Pedagogy of Web-based learning
 - Web-based course delivery systems
 - Design of Distance Learning Environments
 - Cooperative/collaborative learning
 - Corporate training
 - Authoring tools
 - Virtual universities
- Semantic Web
 - Ontologies
 - RDF, DAML, OIL, Topic maps
 - Automated reasoning
 - Annotation