

XHTML Introduction

Fajiang Yu, fjyu@whu.edu.cn

School of Computer, Wuhan University

2017.3



Agenda

- 1 What is XHTML
- 2 XHTML 1.0 Document Conformance
- 3 Main Differences with HTML 4



Agenda

- 1 What is XHTML
- 2 XHTML 1.0 Document Conformance
- 3 Main Differences with HTML 4



What is XHTML

HTML is probably the most successful document markup language in the world.

But XML was introduced. (Why ?)

Why is XHTML (Extensible HTML) needed?

- We could also take advantage of the redesign to clean up some of the more untidy parts of HTML,
- and add some new needed functionality, like better forms.



A brief history of XHTML

XHTML 1.0

XHTML Basic

XHTML Modularization

XHTML 1.1 - Module-based XHTML

XHTML-Print

XML Events (renamed from "XHTML Events")

Other Drafts: XHTML 2.0, etc.

2010-12-17, The XHTML2 Working Group is closed.



XHTML 1.0 is the first document type in the XHTML family. It is a reformulation of the three HTML 4 document types as applications of XML 1.0.



Agenda

- 1 What is XHTML
- 2 XHTML 1.0 Document Conformance
- 3 Main Differences with HTML 4



XHTML 1.0 Document Conformance

- 1) It must conform to the constraints expressed in one of the three DTDs.
 - i) xhtml1-strict.dtd
 - ii) xhtml1-transitional.dtd
 - iii) xhtml1-frameset.dtd
- 2) The root element of the document must be **html**.
- 3) The root element of the document must contain an **xmlns** declaration for the XHTML namespace. The namespace for XHTML is defined to be **http://www.w3.org/1999/xhtml** e.g.

```
<html xmlns="http://www.w3.org/1999/xhtml">
```

- 4) There must be a DOCTYPE declaration in the document prior to the root element.
- 5) The DTD subset must not be used to override any parameter entities in the DTD.



XHTML 1.0 Document DOCTYPE declaration

```
<!DOCTYPE html  
  PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"  
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

```
<!DOCTYPE html  
  PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"  
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
```

```
<!DOCTYPE html  
  PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"  
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">
```



XHTML 1.0 Document XML declaration

An XML declaration is **not required** in all XML documents; however XHTML document authors are **strongly encouraged** to use XML declarations in all their documents.

Such a declaration **is required** when the character encoding of the document is other than the default UTF-8 or UTF-16 and no encoding was determined by a higher-level protocol.



XHTML 1.0 Document Structure

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html
    PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-Transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
    xml:lang="en" lang="en">
  <head>
    <title>XHTML 1.0 Document</title>
  </head>
  <body>
    <p>XHTML 1.0 Document</p>
  </body>
</html>
```



Agenda

- 1 What is XHTML
- 2 XHTML 1.0 Document Conformance
- 3 Main Differences with HTML 4



Main Differences with HTML 4

- 1) Documents must be well-formed
CORRECT: nested elements.
INCORRECT: overlapping elements.
- 2) Element and attribute names must be in lower case
- 3) For non-empty elements, end tags are required
CORRECT: terminated elements
`<p>here is a paragraph.</p>`
`<p>here is another paragraph.</p>`
INCORRECT: unterminated elements
`<p>here is a paragraph.`
`<p>here is another paragraph.`



Main Differences with HTML 4

4) Attribute values must always be quoted

5) Attribute Minimization

CORRECT: unminimized attributes

```
<dl compact="compact">
```

INCORRECT: minimized attributes

```
<dl compact>
```

6) Empty Elements

CORRECT: terminated empty elements

```
<br/><hr/>
```

INCORRECT: unterminated empty elements

```
<br><hr>
```



Thank You!
Any Questions?

