

Introduction à xhtml

François Médard, Jean-Marie Favreau

LASMEA (UMR 6602 UBP/CNRS), LIMOS (UMR 6158 UBP/CNRS)

31 mars 2008

- 1 Introduction
 - Un peu d'histoire...
 - Standardisation
 - xhtml, un dialecte xml
- 2 Le langage xhtml
 - Structure générale d'un document
 - Les balises fondamentales
- 3 Un jeu de balises complet
 - Quelques balises...
- 4 Accessibilité
 - Pourquoi ?
 - Comment ?
- 5 Conclusion

xhtml : l'esperanto du web



Le *World Wide Web*, basé sur internet, lieu de partage d'informations :

- **Documents** structurés par **liens hypertextes**
- Contient des documents **multimédia** : textes, images, vidéos, etc
- On peut **consulter**, mais aussi **créer**
- Nécessité d'un **format commun**

Un peu d'histoire...

1 Introduction

- Un peu d'histoire...
- Standardisation
- xhtml, un dialecte xml

2 Le langage xhtml

- Structure générale d'un document
- Les balises fondamentales

3 Un jeu de balises complet

- Quelques balises...

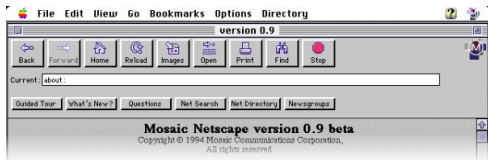
4 Accessibilité

- Pourquoi ?
- Comment ?

5 Conclusion

Un peu d'histoire...

Au début était le html



L'histoire du WWW et du HTML (HyperText Markup Language) :

- 1991 : Tim Berners-Lee (CERN) **annonce** publiquement le web sur Usenet
- 1993 : le navigateur NCSA Mosaic **étend** les idées originales
- 1994 : le navigateur Netscape **démocratise** le HTML
- 1995-1997 : le W3C travaille sur une **spécification du HTML**

1 Introduction

- Un peu d'histoire...
- **Standardisation**
- xhtml, un dialecte xml

2 Le langage xhtml

- Structure générale d'un document
- Les balises fondamentales

3 Un jeu de balises complet

- Quelques balises...

4 Accessibilité

- Pourquoi ?
- Comment ?

5 Conclusion

Le W3C

World Wide Web Consortium fondé en octobre 1994

- Fondé par Tim Berners-Lee
- Objectif : promouvoir la **compatibilité** des technologies du WWW
- Émet des **recommandations** pour le HTML, XHTML, CSS, etc
- HTML :
 - 1995 : RFC 1866 décrivant **HTML 2.0**
 - 1997 : spécification de **HTML 3.2** (janvier), **HTML 4.0** (décembre)
 - 1999 : en décembre, dernière spécification de HTML (**HTML 4.01**)
- XHTML : clarification du format, conjointement avec XML

Vers un format de qualité

Objectifs du XHTML (*extensible* html) :

- **Standardiser** très précisément les pages web
- En faciliter la **lecture automatique**
- Profiter des **outils disponibles** pour le xml : génération et transformation automatique, extraction d'informations, etc.

xhtml, un dialecte xml

1 Introduction

- Un peu d'histoire...
- Standardisation
- **xhtml, un dialecte xml**

2 Le langage xhtml

- Structure générale d'un document
- Les balises fondamentales

3 Un jeu de balises complet

- Quelques balises...

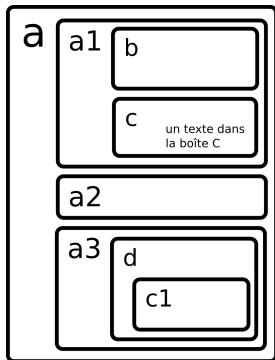
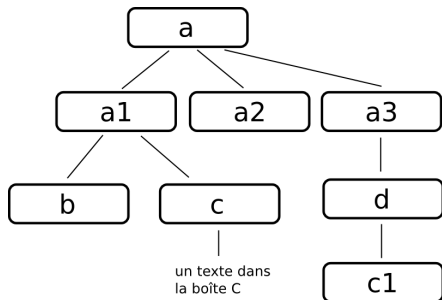
4 Accessibilité

- Pourquoi ?
- Comment ?

5 Conclusion

Le xml en quelques mots

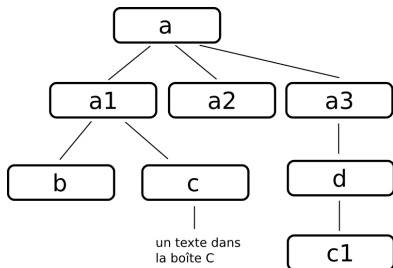
Objectif d'xml : représentation textuelle d'arbres



xhtml, un dialecte xml

Le xml en quelques mots (2)

De l'arbre au texte :



Vocabulaire : racine, nœud (balise)

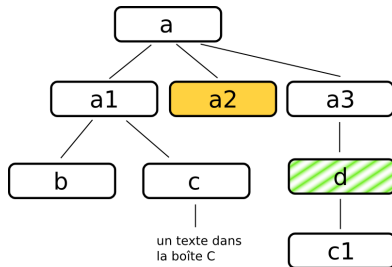
Version avec balises :

```

<?xml version="1.0" encoding="utf-8"?>
<a>
  <a1>
    <b />
    <c>un texte dans la
      boîte C</c>
  </a1>
  <a2 />
  <a3>
    <d>
      <c1 />
    </d>
  </a3>
</a>
  
```

Le xml en quelques mots (3)

De l'arbre au texte :



Vocabulaire : attributs

Version avec balises :

```

<?xml version="1.0" encoding="utf-8"?>
<a>
  <a1>
    <b />
    <c>un texte dans la
      boîte C</c>
  </a1>
  <a2 couleur="jaune" />
  <a3>
    <d fond="rayé">
      <c1 />
    </d>
  </a3>
</a>
  
```

xml : un langage à part entière

- Le langage xml utilise des caractères : <, >, ", '
- Nécessité de les coder :
 - < : <
 - > : >
 - " : "
 - ' : '
 - Et donc & : &
- Existence de commentaires : <!-- commentaire -->

Le xhtml : un dialecte xml

Le xhtml est un **dialecte** du xml :

- Utilisation de la **syntaxe xml**
- Définition de **balises spécifiques**
- Un document est un arbre

Structure générale d'un document

- 1 Introduction
 - Un peu d'histoire...
 - Standardisation
 - xhtml, un dialecte xml
- 2 Le langage xhtml
 - Structure générale d'un document
 - Les balises fondamentales
- 3 Un jeu de balises complet
 - Quelques balises...
- 4 Accessibilité
 - Pourquoi ?
 - Comment ?
- 5 Conclusion

Séparer contenu et apparence

Le html décrivait à la fois **apparence et contenu**. Le xhtml **sépare** les deux :

- **Changer l'apparence** suivant le média (écran, imprimante, pda, ...)
- Faciliter l'**accessibilité**
- Simplifier la maintenance

Le **xhtml** pour le contenu, le **css** pour l'apparence

Données et méta-données

Document xhtml composé de deux parties :

- **Les méta-données** : date de création, auteur, langue, mots-clé, etc.
- **Les données** : contenu de la page

Sont précédées d'une **description du langage utilisé**

Données et méta-données (aperçu)

Structure :

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
  xml:lang="fr" lang="fr">
  <head>
    <title>Page</title>
  </head>
  <body>
    Le contenu ici...
  </body>
</html>
```

L'entête (1)

L'entête (head) devrait contenir :

- L'**encodage** de la page :

```
<meta http-equiv="Content-Type" content="text/html;  
charset=utf-8" />
```

- Les **mots-clés** :

```
<meta name="keywords" content="xhtml, présentation" />
```

- Le **titre de la page** : `<title>L'entête</title>`

L'entête (2)

L'entête (head) peut contenir :

- L'**auteur** :

```
<meta name="author" content="F Médard et Jm Favreau" />
```

- Des informations spécifiques comme le **copyright** :

```
<link rel="copyright"  
href="http://www.gnu.org/copyleft/fdl.html" />
```

- La **feuille css** utilisée :

```
<link href="style.css" rel="stylesheet" type="text/css"  
title="Feuille de style" />
```

- ...

Le corps

Le corps (body) contient tout ce qui sera affiché :

- Les menus
- Le texte
- Les images
- Les liens
- ...

Les balises fondamentales

- 1 Introduction
 - Un peu d'histoire...
 - Standardisation
 - xhtml, un dialecte xml
- 2 Le langage xhtml
 - Structure générale d'un document
 - Les balises fondamentales
- 3 Un jeu de balises complet
 - Quelques balises...
- 4 Accessibilité
 - Pourquoi ?
 - Comment ?
- 5 Conclusion

Titres et paragraphes

Un document est composé :

- Des **titres**, de différent niveau :
 - `<h1>...</h1>`
 - `<h2>...</h2>`
 - ...
 - `<h6>...</h6>`
- Des **paragraphes** : `<p>...</p>`

Titres et paragraphes (exemple)

Structure :

```
<body>
  <h1>Mon titre</h1>
  <h2>Première partie</h2>
  <p>Fusce vulputate lacus
at ipsum. Nam condimentum.
Nulla ut mauris.</p>
  <h2>Suite</h2>
  <h3>Du contenu...</h3>
  <p>Curabitur adipiscing,
mauris non dictum aliquam.
</p>
</body>
```

Mon titre

Première partie

Fusce vulputate lacus at ipsum. Nam condimentum. Null

Suite

Du contenu...

Curabitur adipiscing, mauris non dictum aliquam.

Balises de texte

Se placent dans le texte :

- Avant le xhtml : gras (), italique (<i></i>), police (...)
- Avec le xhtml : **sémantique** plutôt que mise en forme
 - Emphase : ...
 - Emphase forte : ...
 - Balise de style : ... (exemple : ...)

Balises de texte (exemple)

Structure :

```
<body>
  <h1>Mon titre</h1>
  <h2>Première partie</h2>
  <p><strong>Fusce</strong>
vulputate lacus at
<em>ipsum</em>.Nam
condimentum. Nulla ut
mauris.</p>
<h2>Suite</h2>
<h3>Du contenu...</h3>
<p><em>C</em>urabitur
adipiscing, mauris
non dictum aliquam.
</p>
</body>
```

Mon titre

Première partie

Fusce vulputate lacus at *ipsum*. Nam condimentum.

Suite

Du contenu...

Curabitur adipiscing, mauris non dictum aliquam.

Les listes

Les listes :

- Liste **simple** : `...` ou liste **ordonnée** : `...`
- Contient des éléments : `...`
- Chaque élément peut contenir du texte

Les listes (exemple)

Structure :

```
<body>
  <ul>
    <li>Élément 1</li>
    <li>Élément 2</li>
    <li>Élément 3</li>
  </ul>
  <ol>
    <li>Élément 1</li>
    <li>Élément 2</li>
    <li>Élément 3</li>
  </ol>
</body>
```

- Élément 1
- Élément 2
- Élément 3

1. Élément 1
2. Élément 2
3. Élément 3

Les boîtes

Les boîtes `<div>...</div>` :

- **Contiennent** les autres balises
- **Structurent** l'ensemble (menu, contenu, entête, etc)
- Par défaut, **pas d'apparence** (comme les ``)
- Comportement modifié par la **css** associée

Les boîtes (exemple)

Structure :

```
<body>
  <div class="entete">
    <h1>Entete</h1>
  </div>
  <div class="contenu">
    <p>Ceci est le contenu</p>
  </div>
  <div class="pied">
    <p>Ceci est le pied</p>
  </div>
</body>
```

Entete

Ceci est le contenu

Ceci est le pied

Quelques balises...

- 1 Introduction
 - Un peu d'histoire...
 - Standardisation
 - xhtml, un dialecte xml
- 2 Le langage xhtml
 - Structure générale d'un document
 - Les balises fondamentales
- 3 Un jeu de balises complet
 - Quelques balises...
- 4 Accessibilité
 - Pourquoi ?
 - Comment ?
- 5 Conclusion

Les liens externes

C'est la base de l'*hypertexte*, il comporte :

- Un texte apparent (lien vers mes photos)
- Une adresse cible (`http://monsitperso.com/photos`)
- Et d'autres informations optionnelles (langue, titre, etc)

```
<a href="http://monsitperso.com/photos">lien vers mes photos</a>
```

Notion de lien **relatif** et **absolu**

Quelques balises...

Les liens externes (exemple)

Structure :

```
<body>
  <p>Un lien
    <a href="autre.html">relatif</a>.
  </p>
  <p>Un lien
    <a href="http://gnu.org">absolu</a>.</p>
  <a href="other.html" hreflang="en">en
    anglais</a>.</p>
  <p>Un lien
    <a href="other.html" title="lien
      avec titre">avec titre</a>.</p>
</body>
```

Un lien [relatif](#).Un lien [absolu](#).Un lien [en anglais](#).Un lien [avec titre](#).lien avec
titre

Quelques balises...

Les liens internes

Reprend la **même syntaxe**, mais **interne à la page**

- Notion d'ancre : attribut `id` (*identifiant*)
- Le lien interne comporte un **texte** et une **ancre cible**

Attention : un identifiant est unique

Quelques balises...

Les liens internes (exemple)

Structure :

```
<body>
  <p id="premier">Ceci est le premier
  paragraphe.</p>
  <p>...</p>
  <p>Retourner au
  <a href="#premier">premier
  paragraphe</a>.</p>
</body>
```

Ceci est le premier paragraphe.

...

Retourner au [premier paragraphe](#).

Indice, exposant

Le style css permettra de jouer sur la taille, la forme, mais pas le **positionnement en indice ou exposant**.

- Texte en exposant : balise `^{...}`
- Texte en indice : balise `_{...}`

En général, les navigateurs diminuent la taille de la police

Quelques balises...

Indice, exposant (exemple)

Structure :

```
<body>
  <p>Ceci est un texte<sub>avec
  indices</sub>.</p>
  <p>On utilise aussi des
  exposants<sup>comme ici</sup>.</p>
</body>
```

Ceci est un texte_{avec indices}.

On utilise aussi des exposants^{comme ici}.

Les images

Au delà du texte : **inclusion d'images**

- L'image n'est **pas contenue** dans le html
- La balise `` indique au navigateur **où la trouver**
- Formats d'images : jpg, png, gif
- Ajout d'un texte alternatif

Exemple : ``

Quelques balises...

Les images (exemple)

Structure :

```
<body>  
  <p>  
  Image d'un GNU</p>  
</body>
```



Image d'un GNU

Quelques balises...

Les tableaux

Ils servent à structurer présenter des **données structurées**.

- `<table>...</table>`
- Description **par ligne** : `<tr>...</tr>`
- Dans chaque ligne, des **cellules** : `<td>...</td>`
- Existence de **cellules entête** : `<th>...</th>`
- Ajout d'une **légende** : `<caption>...</caption>`
- Ajout d'un **résumé** : attribut `summary="..."`

Remarque : Ne sert pas à structurer une page

Quelques balises...

Les tableaux (exemple)

Structure :

```
<body>
  <table summary="Petit exemple de
    ce qu'on peut faire">
    <caption>Exemple</caption>
    <tr>
      <th>Colonne 1</th>
      <th>Colonne 2</th>
    </tr>
    <tr>
      <td>0.5</td>
      <td>0.2</td>
    </tr>
    <tr><td>1.5</td>
      <td>-0.2</td></tr>
  </table>
</body>
```

Exemple

Colonne 1	Colonne 2
0.5	0.2
1.5	-0.2

Références et citations

Balises sémantiques :

- Une référence est exprimée par la balise `<cite>...</cite>`
- Une citation est exprimée par la balise `<blockquote>...</blockquote>`

Avec différents attributs porteurs d'information : `cite="..."` pour décrire la référence, `lang="..."` pour décrire la langue (dans le cas `<blockquote>`)

Quelques balises...

Références et citations (exemple)

Structure :

```
<body>
<p><cite>Blaise Pascal</cite> a
écrit&nbsp;:</cite></p>
<blockquote lang="fr"
cite="Pensées, Blaise Pascal">
  <p>Deux excès : Exclure la raison,
  n'admettre que la
  raison.</p></blockquote>
</body>
```

Blaise Pascal a écrit :

Deux excès : Exclure la raison,

Pourquoi ?

- 1 Introduction
 - Un peu d'histoire...
 - Standardisation
 - xhtml, un dialecte xml
- 2 Le langage xhtml
 - Structure générale d'un document
 - Les balises fondamentales
- 3 Un jeu de balises complet
 - Quelques balises...
- 4 **Accessibilité**
 - **Pourquoi ?**
 - Comment ?
- 5 Conclusion

L'accessibilité pourquoi ?

L'accessibilité numérique :

- Pour les **personnes handicapées** (matériels et logiciels adaptés)
- Pour **tous les publics** (connections bas-débit, personnes peu compétentes en informatique, configurations spécifiques)
- Pour **rationaliser** le contenu (ergonomie, amélioration de la forme)
- Améliorer l'**interopérabilité** (respect des standards)

Comment ?

- 1 Introduction
 - Un peu d'histoire...
 - Standardisation
 - xhtml, un dialecte xml
- 2 Le langage xhtml
 - Structure générale d'un document
 - Les balises fondamentales
- 3 Un jeu de balises complet
 - Quelques balises...
- 4 **Accessibilité**
 - Pourquoi ?
 - **Comment ?**
- 5 Conclusion


Les images

Faciliter la lecture des images :

- Éviter de faire porter une information **uniquement par l'image**
- Remplir l'attribut `alt="..."`
- Si besoin, utiliser `longdesc="..."`
- Les images significatives doivent être dans le xhtml (et non css)

```

```

 Dessin de gnu

Les liens

Faciliter l'utilisation des liens :

- Utiliser des textes **explicites** (éviter le *cliquer ici*)
- Utiliser l'attribut `lang="..."`



L'utilisation des tableaux

Un héritage malheureux du passé :

- Éviter d'utiliser les tableaux pour structurer le document
- Tableaux utiles pour présenter des données

Structurer le contenu

La structure d'une page facilite son accès :

- **Séparer** entête, menu, contenu(s)
- Placer des liens internes d'**accès rapide**
- **Éviter les cadres** pour structurer le site entier
- Penser à **séparer contenu et apparence**

Respecter les recommandations

Le respect des recommandations :

- Assure un **accès identique** (matériels et logiciels variés)
- Évite de s'appuyer sur les **spécificités** de chaque configuration
- Facilite le fonctionnement des logiciels (lecture automatique, etc)
- Permet le traitement automatique

Conclusion

Le xhtml, esperanto du web :

- **Recommandations** pour un langage universel
- Accès au plus grand nombre à l'information
- Grande richesse **sémantique** des balises
- Un peu de pratique : offre des **possibilités infinies**

Bibliographie

- <http://fr.wikipedia.org/wiki/Html>
- <http://openweb.eu.org>
- <http://validator.w3.org>
- <http://www.w3.org>
- <http://msconline.maconstate.edu/tutorials/XHTML/>