



Multimédia

Projet M18 : HTML5

Aykut ACIKEL / kutay@free.fr
Frédéric MADSEN / madsenfr@laposte.net

M2 MIAGE I

2008/2009

Didier COURTAUD / didier.courtaud@cea.fr

Université d'Évry Val d'Essonne / <http://www.univ-evry.fr>

Version du lundi 16 février 2009





Sommaire

1	Sujet	3
2	Cahier des charges	3
2.1	Étapes de travail.....	3
2.2	Répartition des tâches.....	3
2.3	Matériel rendu.....	4
3	Mises en gardes	4
4	Historique	4
4.1	WHATWG vs W3C	5
5	Contexte	5
5.1	HTML vs XHTML.....	6
6	Nouveautés HTML 5.....	6
6.1	Présentation.....	6
7	Description du site	7
7.1	Fonctionnement.....	7
7.2	Développement	8
7.2.1	Contexte général.....	8
7.2.2	DOCTYPE & charset	8
7.2.3	Validateurs.....	9
8	Modifications en HTML 5.....	9
8.1	Ce qui disparaît	9
8.2	Ce qui change	9
9	Référencement	10
10	Point de vue	10
11	Outils	10
12	Sources	11
12.1	Officielles.....	11
12.2	Non-officielles.....	11



1 Sujet

MI8 : Support de HTML5

HTML5 est la nouvelle version du langage HTML proposé par le WHATWG qui reprend des balises et des fonctionnalités de :

- XUL
- SVG
- SMIL

But du projet :

- Étudier HTML5
- Tester le support des différentes fonctionnalités sur :
 - Firefox 3.1
 - Chrome
 - Opéra
 - IE > 7

2 Cahier des charges

Nous reformulons ici le sujet tel que nous l'avons compris en indiquant de plus la manière dont nous allons procéder pour répondre.

2.1 Étapes de travail

Nous allons procéder en 3 étapes :

1. Détail des nouveautés par rapport à la version précédente (HTML4.01).
2. Mise en œuvre de ces nouveautés via un site web de démonstration.
3. Phase de benchmark des navigateurs Firefox, Internet Explorer, Chrome et Opéra (dans leur dernière version) à l'aide de ce site.

2.2 Répartition des tâches

Nous allons nous répartir ces nouveautés et les étudier en détail tout en mettant les pages adéquates sur notre site afin d'illustrer leur intérêt et leur fonctionnement.

Lorsque cette étape sera finalisée nous partagerons les navigateurs à étudier et réaliserons notre étude de compatibilité sur la base du site ainsi créé.



2.3 Matériel rendu

- Le présent rapport enrichi de l'étude des nouveautés (fonctionnement, utilité, implémentation, etc.).
- Un site web illustrant ces nouveautés (<http://mi8.free.fr/>).
- Une matrice récapitulative de la compatibilité des navigateurs étudiés par rapport à ces nouveautés.

3 Mises en gardes

Les spécifications HTML5 du W3C étant encore au stade de brouillon, le contenu de ce rapport est fortement lié à sa date de réalisation qui est de ce fait rappelée au bas de chaque page.

De plus, le « benchmarking » de navigateurs étant aussi fortement lié à leurs versions, celles-ci sont indiquées dans la section notée « Outils » en fin de ce document.

Les versions dites « nightly build » ne sont pas traitées car trop instables. Seules les versions alpha et bêta des navigateurs qui sont des versions qui préfigurent au plus juste les prochaines versions officielles, sont étudiées. Durant le projet, plusieurs brouillons se sont succédés à un rythme très soutenu, indiquant par là même que la spécification est « bien vivante ».

Lorsqu'une nouveauté est considérée comme supportée par un navigateur, il se peut toutefois qu'elle ne le soit pas entièrement.

Par ailleurs, on s'intéresse ici au support natif par les navigateurs, mais il se peut qu'une extension pallie à ce problème. Dans ce cas le navigateur est considéré non compatible, mais la rubrique « Note » précise le cas échéant où trouver le plugin adéquat.

Le site servant de support de présentation, il peut évoluer légèrement jusqu'à la soutenance.

Certaines nouveautés peuvent donc encore apparaître afin d'être le plus possible exhaustif.

Le site n'est donc pas contractuel au moment de la remise de ce rapport.

4 Historique

HTML est en constante évolution depuis son introduction sur Internet au début des années 1990.

Certaines caractéristiques furent introduites dans des spécifications, d'autres par des éditeurs.

À certains égards, les implémentations et les pratiques des éditeurs ont convergé les unes vers les autres et vers les spécifications et les standards, mais à d'autres égards elles continuent à diverger.



HTML 4 devint une Recommandation du W3C en 1997.

Malheureusement cette recommandation n'est pas assez solide pour permettre l'interopérabilité des applications web.

Il en est de même pour XHTML1, qui définit une sérialisation en XML de HTML 4, et pour DOM 2, qui définit les APIs JavaScript pour HTML et XHTML à la fois.

Le brouillon HTML 5 reflète un effort, commencé en 2004, d'étudier les implémentations HTML contemporaines.

En conclusion, HTML n'a pas évolué depuis une dizaine d'années, et XHTML 1 n'est qu'une version XML de HTML 4 apportant ainsi rigueur et compatibilité avec la boîte à outils XML, mais rien de nouveau pour les utilisateurs du web.

4.1 WHATWG vs W3C

- WHATWG : Web Hypertext Application Technology Working Group
- W3C: World Wide Web Consortium

À l'origine le WHATWG a été créé par Apple, Mozilla et Opera suite à un désaccord avec le W3C.

Ce groupe de travail se limite aux technologies qu'il estime implémentables dans les navigateurs Web sur la base des implémentations actuelles, et particulièrement de celles d'Internet Explorer. Il se présente notamment comme une réponse à la lenteur supposée du développement des standards par le W3C et au caractère supposé trop fermé de son processus interne d'élaboration de spécification. Cependant, de nombreux participants à ce projet sont également des membres actifs du W3C, et le nouveau groupe de travail HTML du W3C a adopté en 2007 les propositions du WHATWG comme base de travail d'un futur HTML5.

Un seul éditeur (Ian Hickson) centralise les propositions et fait évoluer la spécification.

HTML 5 n'est pas encore une spécification aboutie, mais de nombreux navigateurs implémentent déjà certains aspects de la spécification.

Le groupe de travail ne prévoit pas que HTML 5 devienne recommandation avant le troisième trimestre 2010.

5 Contexte

Malgré la stagnation du format HTML, le web a considérablement évolué et des besoins sont apparus, ainsi que des pseudo-normes, poussés par l'industrie du web.



Cela ne s'est pas fait sans que l'interopérabilité en pâtisse, ce qui est pourtant primordiale devant le volume grandissant de données échangées sur le net :

- L'utilisation de vidéos sur internet se fait via des plugins propriétaires et il est difficile d'y accéder de manière programmatique.
- Actuellement le drag n' drop (glisser/déposer) n'est accessible que par programmation principalement JavaScript et présente certains dysfonctionnements.

Cette nouvelle mouture de HTML vise à aller de l'avant tout en conservant une rétrocompatibilité.

Le but avoué étant de rendre les applications dynamiques sur le web plus faciles à réaliser (audio/vidéo, graphiques, communication entre documents, accès hors-ligne, formulaires enrichis, etc.).

Certaines applications telles que les forums, wikis, achats en ligne, recherche, e-mail, messagerie instantanée, chat, édition de document en ligne, sites de partage et sites Web 2.0 en général, deviennent courantes et nécessitent un format de document plus adapté.

5.1 HTML vs XHTML

XHTML 2 encore à l'état de « Draft » et qui tarde à sortir a renforcé les mécontentements dans leur volonté d'agir.

D'où la création du WHATWG avec le HTML 5.

De plus, on reproche au XHTML 2 sa mauvaise rétro compatibilité.

Tout ce qu'on peut dire à l'heure actuelle c'est que la spécification HTML 5 est très active contrairement à XHTML 2, et qu'elle est poussée par l'industrie.


6 Nouveautés HTML 5

On considère bien sûr les différences par rapport à la précédente version. Certaines nouveautés sont directement inspirées d'autres formats comme XUL, SVG, SMIL ou encore XForms, et sont intégrées de manière édulcorée dans la norme HTML 5.

6.1 Présentation

Ces nouveautés sont regroupées en 10 catégories :

- Applications : les nouvelles API
- Cross-document messaging : la communication entre documents
- DOM : les nouveautés, ou inclusion dans la norme

- 
- Édition : les possibilités d'édition de document en ligne
 - Formulaire : enfin de vrais formulaires
 - Interactivité : normalisation de pratiques JavaScript
 - JavaScript : bien que faisant parti d'une spécification à part
 - Multimédia : intégration native du phénomène multimédia
 - Sémantique : pas très démonstratif, mais essentiel
 - Stockage : indispensable pour toute application

Cette répartition est discutable, mais a l'avantage de mettre en lumière les points importants d'innovation de cette nouvelle version.

Les nouveautés HTML 5 en font non plus essentiellement un langage de présentation mais aussi un langage d'application, d'où la nécessité de communication, de stockage et d'interactivité.

L'ensemble de ces rubriques est traité directement sur le site :

- <http://mi8.free.fr>.

7 Description du site

L'intérêt de ce site est de mettre en application directement cette nouvelle norme et de faciliter le test du support par différent navigateur maintenant et dans le futur.

Les exemples de codes ont volontairement été choisis de façon à être minimalistes et à faire apparaître la nouveauté dans toute sa simplicité (quand cela était possible).

Ainsi cette étude n'est pas simplement le fait de lectures, aussi intéressantes soient-elles, mais bien d'une réelle expérience de HTML 5.

7.1 Fonctionnement

Une page d'accueil présente le but du site et ses points clés.

Vous y trouverez une matrice récapitulative des nouveautés et de leurs supports sur les différents navigateurs.

Cette matrice à aussi une fonction de menu pour accéder aux pages traitant de chaque nouveauté.

Chaque page comporte sa date de dernière modification, ce qui est un élément essentiel compte tenu de l'évolution du sujet traité.

Chaque page comporte aussi l'indication du navigateur en cours d'utilisation, ce qui peut influencer sur l'interprétation des codes HTML 5.

Chaque page comporte un lien sur le validateur HTML du W3C et confirme qu'à la date de dernière modification cette page était valide.

Un bouton « Code source » dans le pied de page invite à consulter le code HTML 5.



Pour info seuls Chrome et Firefox offrent cette possibilité nativement, mais cette fonctionnalité est émulée en JavaScript pour les autres navigateurs, ce qui fait que le code n'est peut-être pas aussi bien présenté (indentation, colorisation syntaxique, etc.).

Enfin, on trouve des liens sur le rapport actuel, une archive permettant de récupérer tous les fichiers et un moyen de contacter les auteurs.

7.2 Développement

7.2.1 Contexte général

De part la nature itérative du site, il est vite apparu évident qu'il fallait normaliser les pages de manière à rendre la navigation plus homogène.

Pour ce faire, une modeste librairie PHP développée pour l'occasion permet de constituer aisément de nouvelles pages.

Cette librairie assure la cohérence de chartre graphique, la présence d'éléments techniques redondants (en-tête et pied de page), ainsi que la cohérence des titres de chaque page de nouveauté avec la matrice de compatibilité de la page d'accueil.

7.2.2 DOCTYPE & charset

Simplification notoire par rapport à HTML 4 et sérialisation XML possible.

HTML 4 :

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

HTML 5 :

```
<!DOCTYPE html>
<html>
<head>
  <meta charset="utf-8"/>
  ...
```




Curieusement le validateur du W3C ne reconnaît pas encore le charset au format HTML 5 et donne un warning, c'est pourquoi nous avons conservé l'ancienne déclaration (HTML 4).

XHTML 5 :

```
<?xml version="1.0" encoding="utf-8"?>
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
  ...
```

7.2.3 Validateurs

Il existe plusieurs validateurs, mais le site utilise celui du W3C, les autres étant apparu lorsque le W3C n'était pas encore impliqué dans HTML 5 et n'ont plus qu'une valeur historique.

À noter que le logo  n'existe pas, et qu'il s'agit d'un logo inventé pour l'occasion.

Malgré tout, toutes nos pages sont valides HTML 5, autant qu'elles peuvent l'être avec une norme en cours de rédaction.

C'est d'ailleurs ce que précise le validateur en conservant toujours un warning indiquant que cette norme est encore en cours de validation.

Il faut donc considérer cette validation à la date de la page.

8 Modifications en HTML 5

8.1 Ce qui disparaît

Ce sont essentiellement les éléments de présentation qui disparaissent. Avantagusement remplacés par CSS ils n'avaient d'ailleurs plus de raison d'être depuis longtemps.

À noter également que les frames sont aussi dépréciées.

Tout cela veut dire que les sites comportant de telles balises vont continuer à être présentés par les navigateurs, mais ne seront pas considérés comme du HTML 5.

8.2 Ce qui change

Certaines balises de présentation sont conservées mais redéfinies.

Elles prennent un aspect sémantique qui ne leur était pas forcément attaché au départ.



De plus, certaines sont conservées à des fins d'accessibilité, qui est un axe important de ce projet, notamment envers les déficients visuels.

Cependant, cette étude ne s'attarde pas sur ces points, relativement moins importants, de redéfinition de balises déjà existantes qui ne constituent pas un axe majeur d'évolution et dont il est difficile de faire la démonstration et qui n'ont pas d'intérêt à être traité en terme de support par ou tel navigateur.

9 Référencement

L'apport sémantique indéniable de HTML 5 permettra une meilleure indexation par les moteurs de recherche et la présentation de résultat plus pertinents.

Mais c'est sans compter les milliards de pages référencées sous des versions HTML antérieures qui viendront « polluer » les résultats sémantiques de HTML 5, ou qui devront être ignorées ou traitées différemment dans certains types de requêtes.

10 Point de vue

HTML 5 a une attitude pragmatique qui consiste en partie à constater ce qui fait cruellement défaut aujourd'hui et que chacun tente de contourner à sa manière et à le proposer de manière native dans HTML de sorte que la réalisation et la maintenance de sites en soit simplifiée.

C'est une intention louable, mais qui risque de « scléroser » le web en le figeant dans son état actuel.

La spécialisation de certaines balises comme celle visant à établir des sections de document plutôt qu'une utilisation générique des div est louable en termes de sémantique mais tend à restreindre pour l'avenir la structure des pages à la vision que nous en avons aujourd'hui.

Ce qui a fait la force d'HTML jusqu'à aujourd'hui malgré toute les critiques qu'on a pu faire, tient au fait que ce langage était générique et permettait toutes les fantaisies.

La grande question est donc peut-on être ouvert et sémantique ?

11 Outils

Voici la liste des outils utilisés dans le cadre de la mise en place du site web servant à « benchmarker » les navigateurs :

- IDE : NetBeans
 - <http://www.netbeans.org>
- Éditeur : PSPad
 - <http://www.pspad.com>



- Client FTP : FileZilla
 - <http://www.filezilla.fr>
- Synchronisation des fichiers locaux et distants : SyncBack
 - <http://www.2brightsparks.com/freeware>
- Navigateurs :
 - Chrome 1.0.154.36
 - <http://www.google.com/chrome>
 - Firefox 3.1 beta 2
 - <http://www.mozilla.com/en-US/firefox/all-beta.html>
 - Internet Explorer 8.0.6001.18241 beta 2
 - <http://www.microsoft.com/france/windows/internet-explorer/beta/welcome/default.html>
 - Opera 10.00 alpha
 - <http://www.opera.com/browser/next>

12 Sources

12.1 Officielles

- WhatWG
 - <http://www.whatwg.org>
- W3C
 - <http://www.w3.org>

12.2 Non-officielles

- Mozilla
 - <https://developer.mozilla.org/Fr>