

Du HTML au XHTML en 10 points clés

1. Pour être valide, un document XHTML doit débuter par les déclarations suivantes (voir le pdf sur les DTD).

```
<?xml version="1.0"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "DTD/xhtml1-transitional.dtd" >
<html xmlns="http://www.w3.org/1999/xhtml">
```



2. Les balises <head> et <title> sont désormais obligatoires.

3. Les balises et leurs attributs doivent être en minuscule.

Incorrect: <DIV ALIGN="Center">Titre</DIV>

Correct: <div align="Center">Titre</div>

4. La valeur d'un attribut doit être entre guillemets.

Incorrect: <table border=0 cellpadding=4 cellspacing=0>

Correct: <table border="0" cellpadding="4" cellspacing="0">

5. Tous les attributs doivent avoir une valeur (pas de minimisation possible).

Incorrect: <td nowrap>Cellule 1</td>

Correct: <td nowrap="nowrap">Cellule 1</td>

6. Toutes les balises ouvertes doivent être refermées ou se terminer par un slash ">".

Incorrect: <p>

Incorrect:

Correct: <p></p>

Correct:

7. Les balises doivent être imbriquées correctement.

Incorrect: Some text

Correct: Some text

8. Les éléments "en-ligne" ne doivent pas contenir des éléments de blocs.

Incorrect: <table><tr><td>Lien</td></tr></table>

Correct: <table><tr><td>Lien</td></tr></table>

NB: en-ligne signifie que la balise n'entraîne pas de retour à la ligne (a, img, input, strong...) alors que les éléments de blocs sont toujours formatés avec une césure à la ligne avant et après le blocs (p, h1, table, ol...)

9. Certaines balises possèdent désormais des attributs obligatoires.

Par exemple la balise doit avoir un attribut "alt".

Ou encore la balise <script> doit posséder l'attribut "type".

10. Les signes (<) et (&) sont significatifs dans un document XHTML, à ne pas employer dans le texte, les scripts et les styles

Incorrect: <p>Maider & Jean-Marc sont partis en vacances.</p>

Correct: <p>Maider & Jean-Marc sont partis en vacances.</p>

Incorrect: <script><!-- contenu du script --></script>

Correct: <script><![CDATA [contenu du script]]></script>

NB: Pour ne pas avoir de problème de traduction; il est souvent plus simple d'externaliser les styles et les scripts.

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

```
<script language="javascript" type="text/javascript" src="menu.js">
```

!DOCTYPE et DTD

La déclaration <!DOCTYPE>, situé avant la balise d'ouverture <HTML>, permet de spécifier la version de la norme utilisée dans le document (HTML 3.2, HTML 4.01, XHTML 2.0...).

C'est un Identificateur Publique de Texte (Public Text Identifier ou PTI).



Les normes évoluant, c'est une indication précieuse pour les navigateurs ou périphériques décodant la page même si de nombreux navigateurs acceptent son omission.

Découpage d'une entrée DOCTYPE type :

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML//EN">
<!DOCTYPE [Balise parente] [Validité] "[Enregistrement]// [Organisation]// [Type] [Etiquette]// [Langue]">
```

[Balise parente] - Indique la balise parente déclarée dans la DTD, pour HTML et XML c'est <html>

[Validité] - Indique si le DOCTYPE est public (PUBLIC) ou issu d'une ressource système (SYSTEM) comme un fichier local ou une URL. Les DTD HTML sont identifiés comme PUBLIC.

[Enregistrement] - Le signe "+" indique que l'organisation qui suit est enregistré à l'ISO. L'IETF et le W3C n'étant pas enregistrés, c'est le signe "-" qui est utilisé.

[Organisation] - Indique l'entité ou l'organisation responsable pour la création et la maintenance du DTD. Pour les DOCTYPE HTML et XHTML, ce sont le IETF et le W3C.

[Type] - C'est le "Public Text Class" - type d'objet référencé. Il y a plusieurs mots-clés possible mais dans le cas d'HTML et de XHTML, c'est DTD (Définition Type de Document).

[Etiquette] - C'est la "Public Text Description" - nom unique permettant d'identifier le type de document.

[Langue] - C'est le "Public Text Language", le code de la langue en ISO 639 (voir pdf sur les langues) utilisé pour la création d'objet. Les DTD HTML et XHTML sont toujours en anglais (EN).

Exemple d'entrée DOCTYPE HTML et XHTML	Signification
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">	HTML version 2.0
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 3.2//EN/">	HTML version 3.2
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN">	HTML version 4.0
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN">	HTML version 4.01
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0//EN">	XHTML version 1.0

Bien que cela ne soit pas obligatoire, on peut valider une entrée DOCTYPE en donnant l'URL d'une DTD.

Cela a pour conséquence d'uniformiser la présentation des pages d'un site, cette "Déclaration Type de Document" indique les balises utilisables, les attributs obligatoires ou optionnels ou encore le rendu des feuilles de styles.

On peut créer ses propres DTD (comme en XML) mais il existe des déclarations types utilisables selon la nature du document:

Exemple d'entrée DOCTYPE HTML et XHTML avec URL de DTD	Signification
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">	Utilisation stricte de la norme HTML 4.01 avec feuilles de styles
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">	A utiliser pour des navigateurs ne prenant pas en charge les feuilles de styles
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN" "http://www.w3.org/TR/html4/frameset.dtd">	Identique à "Transitional" pour des documents utilisant des cadres (frames)
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">	Utilisation stricte de la norme XHTML 1.0 avec feuilles de styles
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">	A utiliser pour des navigateurs ne prenant pas en charge les feuilles de styles
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">	Identique à "Transitional" pour des documents utilisant des cadres (frames)

Codes ISO pour la représentation des noms de langues

ISO 639 Code for the representation of the names of languages

Afin d'indiquer la langue d'un texte dans un fichier XHTML, il faut utiliser les attributs "lang" et "xml:lang".

La langue majoritaire du document est spécifiée dans la balise <html> :

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

Pour une portion de texte dans une autre langue, il faut utiliser la balise :

```
<p>Rouge-gorge en Basque se dit <span lang="eu" xml:lang="eu">xenxengorri</span>.</p>
```



Langue	Code	Famille	Langue	Code	Famille	Langue	Code	Famille
Chona	SN	Africaine	Soundanais	SU	Austronésienne	Hindi	HI	Indienne
Haoussa	HA	Africaine	Tagalog	TL	Austronésienne	Marathe	MR	Indienne
Kinyarwanda	RW	Africaine	Tonga	TO	Austronésienne	Népalais	NE	Indienne
Kirundi	RN	Africaine	Letton	LV	Balte	Ourdou	UR	Indienne
Lingala	LN	Africaine	Lithuanien	LT	Balte	Oriya	OR	Indienne
Sango	SG	Africaine				Pendjabi	PA	Indienne
Setswana	TN	Africaine	Basque	EU	Basque	Sanskrit	SA	Indienne
Siswati	SS	Africaine				Sindhi	SD	Indienne
Sotho du sud	ST	Africaine	Abkhaze	AB	Caucasienne			
Swahili	SW	Africaine	Géorgien	KA	Caucasienne	Albanais	SQ	Indo-européenne
Tsonga	TS	Africaine				Arménien	HY	Indo-européenne
Twi	TW	Africaine	Breton	BR	Celtique			
Wolof	WO	Africaine	Cornique	KW	Celtique	Espéranto	EO	Internationale
Xhosa	XH	Africaine	Gaélique (ecossais)	GD	Celtique	Interlingua	IA	Internationale
Yorouba	YO	Africaine	Gallois	CY	Celtique	Interlingue	IE	Internationale
Zoulou	ZU	Africaine	Irlandais	GA	Celtique	Volapuk	VO	Internationale
Azéris	AZ	Altaïque	Afan (oromo)	OM	Chamitique	Afghan (Pachto)	PS	Iranienne
Bachkir	BA	Altaïque	Afar	AA	Chamitique	Kurde	KU	Iranienne
Kazakh	KK	Altaïque	Somali	SO	Chamitique	Persan (farsi)	FA	Iranienne
Kirghiz	KY	Altaïque				Tadjik	TG	Iranienne
Mongol	MN	Altaïque	Kannada	KN	Dravidienne			
Ouïgour	UG	Altaïque	Malayalam	ML	Dravidienne	Grec	EL	Grecque
Ouzbek	UZ	Altaïque	Tamoul	TA	Dravidienne	Latin	LA	Italique
Tatar	TT	Altaïque	Télougou	TE	Dravidienne			
Turc	TR	Altaïque				Catalan	CA	Romane
Turkmène	TK	Altaïque	Groenlandais	KL	Eskimo	Corse	CO	Romane
			Inuktitut	IE	Eskimo	Espagnol	ES	Romane
			Inupiak	IK	Eskimo	Français	FR	Romane
Aymara	AY	Amérindienne				Galicien	GL	Romane
Guarani	GN	Amérindienne	Estonien	ET	Finno-ougrienne	Italien	IT	Romane
Quechua	QU	Amérindienne	Finnois	FI	Finno-ougrienne	Moldave	MO	Romane
			Hongrois	HU	Finno-ougrienne	Occitan	OC	Romane
Bouthani	DZ	Asiatique	Sami du nord	SE	Finno-ougrienne	Portugais	PT	Romane
Birman	MY	Asiatique				Romanche	RM	Romane
Chinois	ZH	Asiatique				Roumain	RO	Romane
Coréen	KO	Asiatique	Afrikaans	AF	Germanique			
Japonais	JA	Asiatique	Allemand	DE	Germanique	Amharique	AM	Sémitique
Khmer	KM	Asiatique	Anglais	EN	Germanique	Arabe	AR	Sémitique
Laotien	LO	Asiatique	Danois	DA	Germanique	Hébreu	HE	Sémitique
Thaï (siamois)	TH	Asiatique	Féroïen	FO	Germanique	Maltais	MT	Sémitique
Tibétain	BO	Asiatique	Frison	FY	Germanique	Tigrinia	TI	Sémitique
Viétnamien	VI	Asiatique	Islandais	IS	Germanique			
Zhuang	ZA	Asiatique	Luxembourgeois	LT	Germanique			
			Néerlandais	NL	Germanique	Bulgare	BG	Slave
Bêche de mer	BI	Austronésienne	Norvégien	NO	Germanique	Biélorusse	BE	Slave
Bislama	BI	Austronésienne	Suédois	SV	Germanique	Croate	HR	Slave
Fidjien	FJ	Austronésienne	Yiddish	YI	Germanique	Tchèque	CS	Slave
Indonésien	ID	Austronésienne				Macédonien	MK	Slave
Javanais	JV	Austronésienne	Assamais	AS	Indienne	Polonais	PL	Slave
Nauruan	NA	Austronésienne	Bengali	BN	Indienne	Russe	RU	Slave
Malgache	MG	Austronésienne	Bihari	BH	Indienne	Serbe	SR	Slave
Malais	MS	Austronésienne	Cachemiri	KS	Indienne	Serbo-croate	SH	Slave
Maori	MI	Austronésienne	Singhalais	SI	Indienne	Slovaque	SK	Slave
Samoan	SM	Austronésienne	Goudjarati	GU	Indienne	Slovène	SL	Slave
						Ukrainien	UK	Slave

en gras les 18 premières langues du monde

Document protégé par la loi sur le droit d'auteur. Copies ou reproductions réservées à l'usage privé ou pédagogique dans un cercle de classe.

Supports en ligne sur <http://universite.online.fr>

Jean-Marc Loisil [universite@online.fr]