



CODAGE VALIDE DES CARACTÈRES WINDOWS ILLÉGAUX EN HTML ET XHTML

- **Profil** : Expert
- **Technologie** : (X)HTML
- **Auteur** : [Laurent Denis \(openwebgroup/bios/laurent_denis\)](http://openwebgroup/bios/laurent_denis)
- **Mise à jour** : 22/05/2008

En bref

Cette page présente les caractères codés de 128 à 159 - propres à Windows, mais invalides selon les spécifications HTML 4+ et XHTML, ainsi que leurs codes de substitution valides.

Cette page présente les caractères codés de `€` à `Ÿ` - propres à Windows, mais invalides selon [les spécifications HTML 4+ et XHTML \(/ressources/specifications/\)](#).

Encodés incorrectement comme dans la première colonne de résultat du tableau ci-dessous, ces caractères ne seront pas restitués correctement sur d'autres plates-formes que Windows (Linux...), où, au mieux, un caractère générique sera affiché (fréquemment un carré vide).

Ils doivent donc être remplacés par leur équivalent en entités numériques ou entités caractères qui, elles, sont valides et seront correctement restituées.

Ces deux solutions ne sont toutefois pas équivalentes :

- aucune des entités caractères ne s'affiche dans Netscape 4 ;
- le résultat obtenu avec les références numériques est indépendant du *Character Encoding* (ISO-8859-1 ou UTF-8), puisque ces codes relèvent du *Character Set* ISO-10646 utilisé en HTML et en XHTML.

caractères invalides et leurs équivalents				
codage invalide	Entité valide	Référence numérique valide	résultat	Remarques
<code>&#128;</code>	<code>&euro;</code>	<code>&#8364;</code>	€	Euro
<code>&#129;</code>				
<code>&#130;</code>	<code>&sbquo;</code>	<code>&#8218;</code>	,	apostrophe anglaise basse
<code>&#131;</code>	<code>&fnof;</code>	<code>&#402;</code>	<i>f</i>	florin, forte musical
<code>&#132;</code>	<code>&bdquo;</code>	<code>&#8222;</code>	„	guillemet anglais bas
<code>&#133;</code>	<code>&hellip;</code>	<code>&#8230;</code>	...	points de suspension
<code>&#134;</code>	<code>&dagger;</code>	<code>&#8224;</code>	†	obèle, dague, croix (renvoi de notes de bas de page)
<code>&#135;</code>	<code>&Dagger;</code>	<code>&#8225;</code>	‡	double croix
<code>&#136;</code>	<code>&circ;</code>	<code>&#710;</code>	^	accent circonflexe
<code>&#137;</code>	<code>&permil;</code>	<code>&#8240;</code>	‰	pour mille
<code>&#138;</code>	<code>&Scaron;</code>	<code>&#352;</code>	Š	S majuscule avec caron (accent circonflexe inversé) utilisé en tchèque
<code>&#139;</code>	<code>&lsaquo;</code>	<code>&#8249;</code>	<	guillemet simple allemand et suisse, parenthèse angulaire ouvrante
<code>&#140;</code>	<code>&OElig;</code>	<code>&#338;</code>	Œ	Ligature o-e majuscule (absente de la norme ISO-8859-1 pour une raison aberrante (http://www.uzine.net/article1785.html) ...)
<code>&#141;</code>				
<code>&#142;</code>		<code>&#381;</code>	Ž	Z majuscule avec caron (accent circonflexe inversé) utilisé en tchèque. Présent dans le <i>Character Encoding</i> ISO-8859-2
<code>&#143;</code>				
<code>&#144;</code>				
<code>&#145;</code>	<code>&lsquo;</code>	<code>&#8216;</code>	‘	guillemet anglais simple ouvrant(utilisé dans les guillemets doubles)
<code>&#146;</code>	<code>&rsquo;</code>	<code>&#8217;</code>	’	guillemet anglais simple fermant(utilisé dans les guillemets doubles)
<code>&#147;</code>	<code>&ldquo;</code>	<code>&#8220;</code>	“	guillemets anglais doubles ouvrants
<code>&#148;</code>	<code>&rdquo;</code>	<code>&#8221;</code>	”	guillemets anglais doubles fermants
<code>&#149;</code>	<code>&bull;</code>	<code>&#8226;</code>	•	boulet, utiliser plutôt des listes à puces

codage invalide	Entité valide	Référence numérique valide	résultat	Remarques
–	–	–	–	tiret demi-cadratin (incise), voir The Trouble With EM 'n EN (http://www.alistapart.com/stories/emen/)
—	—	—	—	tiret cadratin (dialogue), voir The Trouble With EM 'n EN (http://www.alistapart.com/stories/emen/)
˜	˜	˜	~	tilde
™	™	™	™	marque déposée
š	š	š	š	s minuscule avec caron (accent circonflexe inversé) utilisé en tchèque
›	›	›	>	guillemet simple allemand et suisse, parenthèse angulaire fermante
œ	œ	œ	œ	Ligature o-e minscule (absente de la norme ISO-8859-1 pour une raison aberrante (http://www.uzine.net/article1785.html) ...)
				
ž		ž	ž	z minuscule avec caron (accent circonflexe inversé) utilisé en tchèque. Présent dans le <i>Character Encoding</i> ISO-8859-2
Ÿ	Ÿ	Ÿ	ÿ	Y majuscule avec trema, présent en français dans quelques noms propres (http://www.gutenberg.eu.org/pub/GUTenberg/publicationsPDF/25-andre.pdf) (PDF).

Une question, une remarque ? Écrivez à [l'auteur \(openwebgroup/bios/laurent_denis\)](mailto:openwebgroup/bios/laurent_denis).

[\(/articles/caracteres_illegaux/#forum\)](#)