

Le XML et le HTML

Lorsque les techniques de publication sur le Web sont abordées, une comparaison entre le HTML, le XHTML et le XML s'impose. Au contraire de ce qui a déjà été écrit par ailleurs, le XML n'est pas le successeur du HTML. Le XML, le XHTML et le HTML sont des langages distincts !

1. Une seule similitude : le SGML

Le seul point commun entre le HTML et le XML est qu'ils sont issus tous deux de la même "mère", soit le SGML (*Standardized Generalized Markup Language*), qui est le langage de référence en milieu professionnel pour tout ce qui concerne la gestion électronique des documents. Ils sont donc, tous deux, des langages de balises (voir le *Markup Language*). Ils ont également des caractéristiques communes héritées du SGML, qui consistent à transporter sur le Web des données en mode texte (*plain text*), compatibles avec n'importe quelle plateforme logicielle.

2. Les différences entre le HTML et le XML

Le HTML et le XML diffèrent sur de très nombreux points, dont certains ont trait à l'essence même du langage.

- Le XML décrit, structure, échange des données tandis que le HTML a pour vocation d'afficher des données.
- Le XML est un générateur de langages (métalangage). Le HTML est un langage normalisé de publication sur le Web.
- Le XML est extensible. Il permet de créer ses propres balises en fonction des données traitées. En HTML, les balises sont prédéfinies et elles sont donc figées.
- Le XML est un langage strict dont l'écriture doit être rigoureuse. La syntaxe du HTML est devenue très (trop ?) permissive à cause des navigateurs récents.