

Les langages de schéma XML

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 2 Février 2006. Dernière mise à jour le 16 Février 2006

<http://www.bortzmeyer.org/langages-schema-xml.html>

Je viens de finir un exposé sur les langages de schéma pour décrire un langage XML. Les transparents sont disponibles sous une forme adaptée à la présentation (en ligne sur <http://www.bortzmeyer.org/files/xml-schema-langages-presentation.pdf>) ou bien à l'impression (en ligne sur <http://www.bortzmeyer.org/files/xml-schema-langages-impression.pdf>). Le source (en ligne sur <http://www.bortzmeyer.org/files/xml-schema-langages.tex>) est également disponible.

XML permet de définir plusieurs langages spécialisés, comme Docbook, Atom (format de syndication, RFC 4287¹), ou bien EPP ou IRIS, chers à tous les registres. Chacun de ces langages est défini par un schéma et ce schéma peut être écrit dans différents langages :

- le langage originel, DTD (utilisé par le RFC 2629, par Docbook jusqu'à la version 4, l'actuelle),
 - Examplotron <<http://examplotron.org/>>, qui permet de créer un schéma à partir d'un document exemple,
 - les W3C Schemas, souvent appelés XSD (utilisés par Wikipedia, par SOAP, ou bien pour EPP et IRIS - RFC 3981),
 - RelaxNG, le petit dernier (utilisé par OpenDocument, Docbook à partir de la version 5, Atom, RFC-editor, queue, etc).
- Avoir un schéma pour ses documents XML (il faut noter que ce n'est pas obligatoire) permet de les valider et de guider le processus d'édition.

Cet exposé parle donc de :

- Petit rappel XML. - Pourquoi un langage de schéma ? Validation et édition guidée.
- Caractéristiques d'un langage de schéma. Décrire la structure et décrire le contenu des éléments (leur type). Ambiguïté et non-déterminisme.
- Exemple de langage XML qui va être implémenté avec les divers schémas.
- DTD, l'ancêtre et ses limites.
- Examplotron.
- W3C Schemas, le langage d'EPP et d'IRIS.
- RelaxNG, le langage de Docbook et d'Atom.
- La notion de bibliothèque de types.
- Outils de validation avec un schéma : rnv <<http://www.davidashen.net/rnv.html>>, xml-lint <<http://xmlsoft.org/>> ou bien jing <<http://www.thaiopensource.com/relaxng/jing.html>>.
- Outils d'édition guidée avec un schéma : nxml mode <<http://www.thaiopensource.com/nxml-mode/>>.

1. Pour voir le RFC de numéro NNN, <http://www.ietf.org/rfc/rfcNNN.txt>, par exemple <http://www.ietf.org/rfc/rfc4287.txt>