

The logo for Tirème SARL, featuring the company name in white, uppercase letters on a dark blue rectangular background.

Canonical, XML Canonical

Rédaction : Pierre Attar

Cette spécification décrit ce que doit faire un "canoniseur" XML pour produire une représentation physique (la forme canonique) d'un document XML, qui ne change pas quelles que soient les variations syntaxiques de l'entrée XML.

Avec cette spécification, si un document XML est changé par une application, mais si sa forme canonique XML n'a pas changé, le document changé et le document initial peuvent être considérés comme équivalents par les applications les utilisant.

Cette spécification est mineure dans l'ensemble de celles issues du W3C. Elle sera certainement utilisée dans des ingénieries logicielles très spécifiques, dès lors qu'il sera nécessaire de s'intéresser à des versions de documents.

Objectifs

L'objectif de cette spécification est de pouvoir s'appuyer sur une forme XML canonique pour définir l'équivalence de deux documents XML, de façon indépendante des variations syntaxiques qui peuvent avoir été introduites, comme des variations de codage de caractères (on se souviendra que XML accepte différents codages), d'ordre de définition d'attributs, de syntaxe d'écriture des éléments `empty`, etc.

Cette notion d'équivalence est nécessaire, dès lors que l'on introduit, par exemple, des notions de signature de document (XML-Signature) ou dès lors que l'on s'intéresse, dans XPath, à des positions relatives d'objets les uns par rapport aux autres.

Cette spécification est mineure : son utilité n'est définie que pour des environnements applicatifs extrêmement précis. Elle n'a pas de volonté de devenir une des recommandations de base du W3C, comme peuvent l'être XPath ou XSLT.

Principes

Un "canoniseur" peut générer un document XML "bien formé" ; il peut aussi générer une partie de document. Dans ce dernier cas, la spécification de ce qui doit être canonisé est réalisée en utilisant une expression issue du langage XPath.

D'un point de vue fonctionnel, le codage de caractères du document généré est le codage UTF-8. Basée sur le modèle de donnée de XPath, le reste de la spécification s'intéresse surtout à des notions d'ordonnement d'objets (les attributs, par exemple) et à des notions de normalisation d'attributs et de "retour-chariot".

Canonical définit la spécification d'un outil, de ce qu'il doit interpréter dans un document et ce qu'il doit générer. Cette spécification n'a pas pour objectif d'établir une méthode définissant l'équivalence de deux documents XML dont les formes canoniques sont identiques : elle se contente de donner aux applicatifs les moyens de le faire.

Recommandations(s)

■ ■ *XML canonique*

Recommandation, version 1.0, du 15-03-2001

Document sur <http://www.yoyodesign.org/doc/w3c/xml-c14n/index.html>

🇺🇸 *Canonical XML*

Recommandation, version 1.1, du 02-05-2008

Document sur <http://www.w3.org/TR/xml-c14n>

🇺🇸 *XML Canonicalization Requirements*

Note, version 19990605, du 05-06-1999

Document sur <http://www.w3.org/TR/NOTE-xml-canonical-req>

■ ■ *La canonisation XML exclusive*

Recommandation, version 1.0, du 18-07-2002

Document sur <http://www.yoyodesign.org/doc/w3c/xml-exc-c14n/>

🇺🇸 *Exclusive XML Canonicalization*

Recommandation, version 1.0, du 18-07-2002

Document sur <http://www.w3.org/TR/xml-exc-c14n>