### XML canonic

*XML canonic*, definit de *James Clark*, reprezintă un subset al limbajului XML, util pentru testarea procesoarelor (analizoarelor) XML, ca mod de reprezentare a unui document XML analizat. În prezent, în cadrul Consorţiului Web se lucrează la standardizarea specificaţiei XML canonic.

Pentru fiecare document XML bine-formatat există un document XML canonic echivalent cu documentul original.

Regulile de producţie corespunzătoare gramaticii care generează limbaje XML canonice sunt următoarele:

[0] CanonXML ::= Pi\* element Pi\*

[1] element ::= Stag (Datachar | Pi | element)\* Etag

[2] Stag ::= '<' Name Atts '>'

[3] Etag ::= '</' Name '>'

[4] Pi ::= '<?' Name ' ' (((Char - S) Char\*)? - (Char\* '?>' Char\*)) '?>'

[5] Atts ::= (' ' Name '=' '"' Datachar\* '"')\*

[6] Datachar ::= '&amp;' | '&lt;' | '&gt;' | '&quot;'

| '&#9;' | '&#10;' | '&#13;'

| (Char - ('&' | '<' | '>' | '"' | #x9 | #xA | #xD))

[7] Name ::= (conform specificaţiei XML)

[8] Char ::= (conform specificaţiei XML)

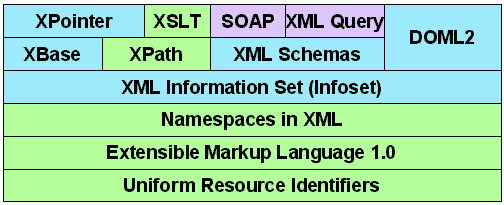
[9] S ::= (conform specificaţiei XML)

Regula [0] indică faptul că un document XML canonic poate începe şi se poate sfărşi cu un număr de zero, una sau mai multe directive de procesare (Pi) şi poate conţine măcar un element, compus dintr-un *tag* de început şi dintr-un *tag* de sfârşit. Între aceste *tag*-uri pot exista date de tip caracter, alte directive de procesare sau elemente (regula [1]). Un *tag* de început este prefixat de caracterul "<" urmat de un nume şi de o mulţime (posibil vidă) de atribute (conform regulilor [2] şi [5]). Un *tag* de sfârşit este prefixat de "/", urmat de numele elementului şi de caracterul ">" (regula [3]). Instrucţiunile de procesare sunt specificate de regula [4], iar atributele de regula [5]. Datele de tip caracter urmează regula [6], iar neterminalii Name, Char şi S sunt aceeaşi din specificaţia XML 1.0.

### 8 Viitor

Limbajul XML oferă o metodă de comunicare inter-umană, nu de interschimb de date între maşini, realizând relaţii între creatori şi beneficiarii documentelor acestora.

XML moşteneşte neutralitatea platformei, limbajului şi libertatea formatelor de date prezente în SGML, dar prin XML Schema câştigă în flexibilitatea specificării formale a tipurilor şi modului de apariţie a elementelor şi atributelor.



##### Nivelele XML

Toate aceste remarci ne dau certitudinea că XML oferă utilizatorilor un standard deschis de marcare a informaţiilor, fără limite în ceea ce priveşte tehnicile de adnotare (extensibilitate asigurată), uşor de folosit (putând conferi suport pentru baze de cunoştinţe), facil de implementat.

Viitorul XML nu va avea implicaţii doar asupra informaticienilor, ci şi asupra lumii economice şi politice. Combinaţia XML cu XSL poate înlocui toate formatele de procesare de texte şi publicare existente în prezent:

* un format unic pentru publicarea pe Web şi tipărituri;
* un format comun pentru stocarea datelor în cadrul diverselor produse software;
* un format unic pentru toate limbile.

Astfel, beneficiile majore ale familiei de limbaje XML pot fi următoarele:

* XML reprezintă un format flexibil de transmitere a informaţiilor pe Web;
* XML este un standard deschis, independent de platforma hardware şi software;
* XML poate genera perspective multiple pentru aceleaşi seturi de date;
* XML oferă suport pentru căutarea inteligentă şi eficientă a datelor pe Web;
* XML permite dezvoltarea de aplicaţii Web reale (concrete);
* XML oferă o scalabilitate (virtual) nelimitată;
* XML furnizează suport pentru colectarea şi integrarea datelor provenite din surse multiple;
* XML dă posibilitatea de înţelegere pentru cititorul uman a documentelor de dimensiuni mari.

Succesul XML poate fi întrevăzut mai ales în lumea spaţiului Web, permiţând crearea de aplicaţii hipermedia complexe, dar şi în procesarea informaţiilor în general.