

XML و ده نکته مهم

دنیای XML مملو از تکنولوژی های دیگر، نظیر XLink, Namespace, DTD, Schema, CSS, XSL, XHTML و ... است. افرادی که قصد آشنائی با XML را دارند، همواره با این سوال مشکل مواجه هستند که از کجا می بایست آغاز نمایند. در این مقاله به بررسی ده نکته مهم بمنظور آشنائی با XML پرداخته تا از این رهگذر مسیری مناسب برای علاقه مندان به فراگیری این تکنولوژی مهم، مشخص گردد.

نکته اول: XML گزینه ای برای ارائه ساختار داده ها

داده های ساختاریافته شامل مواردی نظیر صفحات گسترده، دفترچه های آدرس، پارامترهای قابل پیکربندی و ... می باشند. XML، مجموعه ای از قوانین بمنظور طراحی ساختار داده های مورد نظر است. XML، بعنوان یک زبان برنامه نویسی تلقی نشده و برای استفاده از آن اجباری به داشتن تجربه لازم برنامه نویسی نخواهد بود. با استفاده از XML، کامپیوترها بسادگی قادر به تولید و خواندن اطلاعات و اطمینان از عدم مبهم بودن ساختمان داده ها می باشند. XML، بسط پذیر، مستقل از پلات فورم و مورد حمایت های بین المللی و محلی است. XML، بطور کامل Unicode را حمایت می نماید.

نکته دوم: XML شباهت زیادی به HTML دارد

XML، همانند HTML از تگ ها و خصلت ها استفاده می نماید. در HTML، هر تگ و یا خصلت دارای عملکرد و معنی از قبل مشخص شده ای بوده و یا بکارگیری هر تگ، نحوه نمایش اطلاعات (بین تگ های استفاده شده) برای مرورگر تبیین می گردد. XML، با استفاده از تگ ها، محدوده مشخصی برای داده ها را تعیین و عملاً مسئولیتی در رابطه با تفسیر تگ ها و داده های همراه آنها را نداشته و این عملیات را به برنامه هائی که یک فایل XML را می خوانند، واگذار می نماید. مثلاً در صورتیکه از تگ <p> در یک سند XML استفاده شده باشد، نمی توان این ادعا را داشت که تگ فوق نشاندهنده یک پاراگراف است. ممکن است تگ فوق، نشاندهنده price، parameter و یا یک person باشد.

نکته سوم: XML بصورت متن است.

برنامه هائی که داده های ساختاریافته تولید می نمایند، اغلب داده های خود را بر روی دیسک ذخیره می نمایند. فرمت ذخیره سازی اطلاعات ممکن است بصورت متن و یا باینری باشد. مهمترین مزیت ذخیره سازی داده ها بصورت متن، امکان مشاهده داده ها بدون نیاز به برنامه ای است که اطلاعات را تولید کرده است. حتی می توان اطلاعات فوق را با استفاده از یک ویرایشگر متنی تغییر داد. فرمت های اطلاعاتی مبتنی بر متن، به پیاده کنندگان نرم افزار قابلیت های بیشتری را در خصوص اشکال زدانی برنامه ها نیز می دهد. فایل های XML، مشابه HTML بصورت متن می باشند. قوانین موجود در XML، با صراحت و شفافیت بیشتری ارائه شده است. فراموش کردن یک تگ و یا عدم قرار دادن خصلت بین علامت "" باعث می گردد که یک فایل XML غیرقابل استفاده گردد. در صورتیکه در فایل های HTML مسائل فوق تحمل و اغلب مجاز خواهد بود. اگر یک فایل XML دارای مشکل باشد، برنامه استفاده کننده پس از مواجه شدن با مشکل، عملیات خود را در محل بروز خطا متوقف و یک پیام خطا را ارائه خواهد داد.

نکته چهارم: ظرفیت فایل های XML زیاد است.

با توجه به اینکه فرمت فایل های XML، بصورت متنی بوده و از تگ ها بمنظور مرزبندی داده ها استفاده می شود، این نوع فایل ها معمولاً دارای حجم بیشتری نسبت به فایل هائی با فرمت باینری می باشند. همانگونه که در نکته سوم اشاره گردید، مهمترین مزیت یک فرمت مبتنی بر متن، قابلیت مشاهده آن و مهمترین اشکال، اشغال حجم بیشتری از فضای ذخیره سازی است. از طرف دیگر با توجه به وجود برنامه های فشرده سازی نظیر Zip، امکان فشرده سازی فایل ها بخوبی و با سرعت بالا فراهم می گردد. پروتکل های ارتباطی نظیر پروتکل های مربوط به مودم و یا HTTP/1.1 (پروتکل اصلی در وب) قادر به فشرده سازی داده ها و صرفه جوئی در پهنای باند استفاده شده می باشند.

نکته پنجم: XML خانواده ای از تکنولوژی ها است

XML 1.0، مشخصات، ماهیت و قوانین موجود در رابطه با تگ ها و خصلت ها را تعریف می نماید. خانواده XML، شامل مجموعه ای در حال رشد از سایر ماژول هائی است که سرویس های مفیدی را در خصوص انجام عملیات و خواسته های مهم بر عهده دارند. XLink، روشی استاندارد برای افزودن ابر لینک ها به یک فایل XML را مشخص می نماید. XPointer، گرامری مناسب برای اشاره به بخش های متفاوت یک سند XML است. یک شباهت زیادی به یک URL داشته، با این تفاوت که در مقابل اشاره به سندهائی بر روی وب، به بخش های متفاوت در یک سند XML اشاره می نماید. CSS، زبان style sheet قابل استفاده برای XML و HTML است. XSL، یک زبان پیشرفته برای ارائه sheet style است. تکنولوژی فوق بر اساس XSLT است. XSLT، یک زبان تبدیل بمنظور سازماندهی مجدد، افزودن و یا حذف تگ ها و خصلت ها است. DOM، استاندارد ای با مجموعه ای از توابع برای انجام عملیات بر روی یک سند XML (یا HTML) از طریق یک زبان برنامه نویسی است. XML Schema، به پیاده کنندگان، امکان تعریف دقیق ساختار اطلاعات را ارائه می

نماید. در این رابطه ماژول های متعدد دیگری نیز وجود داشته که در خانواده بزرگ XML دارای جایگاه خاص خود می باشند.

نکته ششم: XML جدید است اما نه خیلی جدید

پیاده سازی XML از سال 1996 آغاز و در فوریه سال 1998 توسط کنسرسیوم وب عرضه گردید. ممکن است ابهامی در خصوص جوان بودن تکنولوژی فوق در ذهن ایجاد گردد. در پاسخ می بایست به این نکته اشاره گردد که تکنولوژی فوق، خیلی هم جدید نیست. قبل از XML، تکنولوژی SGML وجود داشت. پیاده کنندگان در اوایل دهه 80 در مقیاس وسیعی از آن برای پروژه های بزرگ مستند سازی استفاده می نمودند. پیاده سازی HTML از سال 1990 آغاز گردید. طراحان XML، بهترین بخش های SGML را به همراه تجارب موجود در رابطه با HTML، ترکیب و محصولی را تولید نموده اند که قدرتی کمتر از SGML ندارد. SGML اغلب برای مستند سازی فنی استفاده شده و از آن برای سایر موارد مربوط به داده ها بندرت استفاده می گردد. XML دقیقاً "برخلاف وضعیت فوق است".

نکته هفتم: XML باعث هدایت HTML بسمت XHTML می گردد.

یکی از موارد مهم کاربرد XML، تکنولوژی XHTML است. XHTML، گزینه بعدی نسبت به HTML است. XHTML دارای المان های مشابه المان های استفاده شده در HTML است. گرامر استفاده شده در XHTML، بگونه ای تغییر پیدا کرده که منطبق بر قوانین مربوط به XML باشد. سندهای مبتنی بر XML، گرامر خود را از XML به ارث برده و آن را با روش های خاصی محدود می نماید. (مثلاً XHTML امکان استفاده از <p> را داده ولی امکان استفاده از <r> را نخواهد داد) در XHTML به گرامر موجود، مفهوم نیز اضافه شده است. (مثلاً XHTML از تگ <p> برای یک پاراگراف استفاده می نماید نه برای price و یا person)

نکته هشتم: XML ماژولار است.

XML، امکان تعریف یک فرمت جدید سند را با ترکیب و استفاده مجدد از سایر فرمت ها فراهم می نماید. با توجه به اینکه دو فرمت بصورت مستقل از یکدیگر طراحی و پیاده سازی می گردند، ولی ممکن است از المان ها و یا خصلت هایی با اسامی مشابه، استفاده شده باشد. در زمان ترکیب فرمت های فوق، می بایست تدابیر لازم پیش بینی گردد. (مثلاً ممکن است تگ <p> در یک فرمت، مفهوم پاراگراف را داشته در صورتیکه در فرمت دیگر، دارای مفهوم person باشد). بمنظور حل مشکل فوق (وجود اسامی مشابه با مفهوم متفاوت)، namespace ارائه شده است. XSL و RDF نمونه های مناسبی از فرمت های مبتنی بر XML بوده که از namespace استفاده می نمایند.

نکته نهم: XML پایه و اساس RDF و سمنتیک وب است.

RDF (Description FrameWork Resource) یک فرمت متنی مبتنی بر XML بوده که تشریح منابع و برنامه های متادیتا (نظیر یک برنامه موزیک playlist و یا آلبوم عکس) را حمایت می نماید. مثلاً RDF می تواند این امکان را فراهم آورد که یک آلبوم عکس وب را با استفاده از اطلاعات مربوط به لیست تماس شخصی، ایجاد و در ادامه برنامه سرویس گیرنده پست الکترونیکی قادر به ارسال یک پیام بصورت خودکار برای افراد موجود در آلبوم باشد. همانگونه که HTML باعث ترکیب سندها و برنامه های مبتنی بر فرم ها در وب می گردد، RDF با یکپارچگی برنامه ها و آژانس ها، یک وب سمنتیک (وبی که تعامل اطلاعاتی بین اجزای آن کاملاً مشهود و مدون است) ایجاد می نماید.

نکته دهم: حق استفاده از XML رایگان بوده، مستقل از پلات فورم است و بخوبی حمایت می گردد.

با انتخاب XML، برای پیاده سازی یک پروژه از امکانات حمایتی متنوع و گسترده ای در رابطه با ابزارها و مهندسین با تجربه بهره مند خواهید شد. انتخاب XML، مشابه انتخاب SQL برای بانک اطلاعاتی است. در این حالت همچنان نیاز به ایجاد بانک اطلاعاتی، نوشتن برنامه ها و روتین های لازم بمنظور انجام عملیات دلخواه خواهد بود. با توجه به اینکه حق استفاده از XML رایگان بوده می توان نرم افزار خود را با استفاده از آن، بدون پرداخت هزینه ای طراحی نمود.