

ایکس.ام.ال. به عنوان یک زبان ابر داده

نویسنده: مهدی تلخایی

دانشجوی کارشناسی ارشد کتابداری و اطلاع رسانی دانشگاه آزاد همدان و کارشناس بخش مرجع کتابخانه مرکزی دانشگاه اراک
mtlib11@gmail.com

چکیده

ایکس.ام.ال. یک استاندارد صنعتی قدرتمند و انعطاف پذیر برای انتقال اطلاعات ساخت یافته است. این زبان توسط کنسرسیوم شبکه جهانی به عنوان ابزاری برای به اشتراک گذاردن اطلاعات از طریق اینترنت ارائه شده است. مجموعه مشخصات ایکس ام ال توسط یک گروه کاری در کنسرسیوم وب جهانی منتشر شده است و توسط تعدادی از شرکت های پیشرو در صنعت کامپیوتر پشتیبانی می شود.

کلید واژه‌ها

ایکس.ام.ال.، ابر داده، زبان نشانه گذاری

مقدمه

ایکس.ام.ال.^۱ یک استاندارد صنعتی قدرتمند و انعطاف پذیر برای انتقال اطلاعات ساخت یافته است. این زبان توسط کنسرسیوم شبکه جهانی^۲ به عنوان ابزاری برای به اشتراک گذاردن اطلاعات از طریق اینترنت ارائه شده است. این زبان برای کاربردهای غیر اینترنتی نیز مورد توجه بوده است و برای توصیف داده طراحی شده و بر آنچه از نوع داده است تمرکز دارد. ایکس ام ال به طراحان و تولید کنندگان وب امکان می دهد تا برجسب‌هایی^۳ ایجاد نمایند که انعطاف پذیری آنها در سازماندهی و نمایش اطلاعات بیشتر از سیستم‌هایی است که با آن‌ها می‌توان تهیه می‌شوند. مجموعه مشخصات ایکس ام ال توسط یک گروه کاری در کنسرسیوم وب جهانی منتشر شده است و توسط تعدادی از شرکت های پیشرو در صنعت کامپیوتر پشتیبانی می شود (Fitzgerald 2007,129).

ایکس.ام.ال. نیز همانند اچ.تی.ام.ال. نوعی زبان نشانه‌گذاری و زیرمجموعه زبان نشانه‌گذاری عمومی استاندارد است که به صورت ویژه برای ایجاد صفحات وب طراحی شده است. با این تفاوت که ایکس.ام.ال. برخلاف اچ.تی.ام.ال.، که تنها به قالب‌بندی سند یعنی تعیین ساختار و ظاهر آن می‌پردازد، امکان توصیف محتوای صفحات وب را با استفاده از ابر داده فراهم می‌کند و نیز به کاربر اجازه می‌دهد تا بنابر نیاز، علائم نشانه‌گذاری جدیدی تعریف کند. تکامل زبان‌های نشانه‌گذاری و ترویج استفاده از زبان نشانه‌گذاری توسعه‌پذیر (ایکس.ام.ال) در خلق صفحات وب، به تدریج، زمینه ارتقای کیفی خدمات اطلاعاتی وب را فراهم می‌کند.

برنرز، با تکیه بر ایکس.ام.ال. و برخی فناوری‌های دیگر که تحت هدایت کنسرسیوم وب در حال توسعه هستند، ایده وب معنایی^۴ را به عنوان نسل آینده وب مطرح کرد. به زعم وی، وب معنایی "متشکل از داده‌هایی است که به صورت مستقیم یا غیرمستقیم توسط ماشین‌ها قابل پردازش هستند". اکثر صفحات وب کنونی در حقیقت به منظور استفاده انسان - یعنی صرفاً خواننده شدن - خلق می‌شوند. داده‌های وب توسط رایانه‌ها خوانده و شمارش می‌شوند، اما عموماً قابل پردازش نیستند. وب معنایی، که شکل توسعه یافته وب کنونی خواهد بود، با استفاده از ابر داده‌های ساختار بندی شده و تدارک یک بستر مفهومی برای داده‌ها، امکان پردازش داده‌ها را متناسب با نیاز کاربر و توسط نرم‌افزارهای مختلف فراهم می‌کند (Berners-Lee 191,2005).

در محیط وب معنایی، رایانه‌ها می‌توانند از شبکه داده‌های منبع استفاده کرده، آنها را تعبیر، تحلیل، و پردازش کرده و به کاربر ارائه نمایند. در وب معنایی تلاش بر این است تا هر ماده اطلاعاتی (به عنوان مثال یک سند یا عکس) بر اساس موضوع، پدیدآورنده، محل، تاریخ، و سایر خصوصیاتش توصیف شود. این امر موجب تسهیل پردازش و بازیابی داده‌ها شده و امکان خودکارسازی بسیاری از امور را فراهم می‌آورد.

مزایای ایکس.ام.ال.

تمام استانداردهای مهم آموزش الکترونیکی، به جز مکانیزمهای مربوط به بررسی داده‌ها، بر مبنای ایکس.ام.ال. هستند. یادگیری فاصله دار پیشرفته^۵ و کمیته صنعت هوانوردی^۶ تلاشهای زیادی را برای توسعه یک پروتکل برای بررسی داده‌ها صورت دادند. در واقع هدف آنها از انجام این فعالیتها توسعه و پیشرفت ایکس.ام.ال. است. سازمانهای استاندارد کننده، ایکس.ام.ال. را به دلایل متعددی انتخاب کرده اند که برخی از آنها به شرح زیر است (Kurt,2003,95):

۱. این زبان قادر است اطلاعات و ساختار یا مفهوم آنها را انتقال دهد.

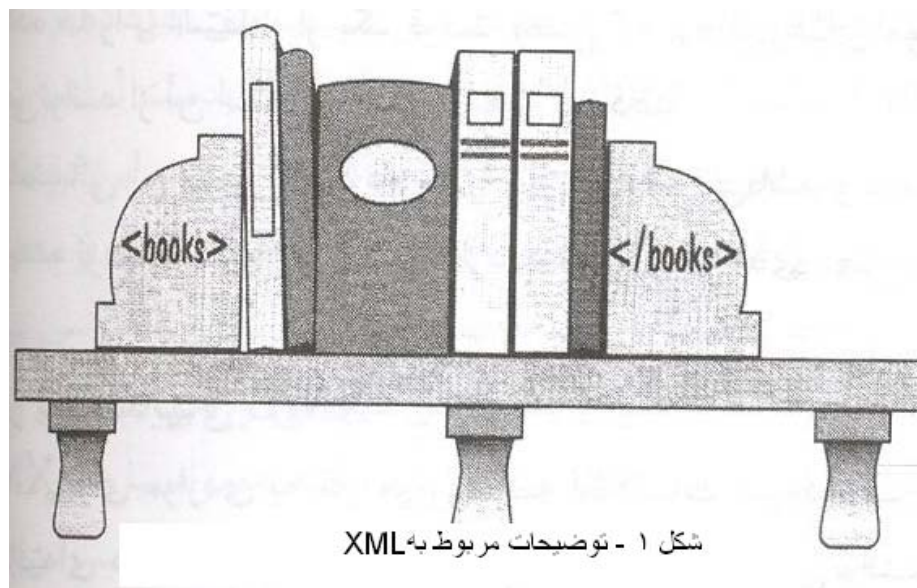
۲. افراد غیر حرفه ای می توانند این زبان نوشته را نوشته، خوانده و درک کنند، بنابراین مختص برنامه نویسان و کامپیوترها نیست.
۳. داده ها را با استفاده از یک فرمت متنی که نرم افزارها موجود می توانند از آنها استفاده کنند، می توان انتقال داد.
۴. پشتیبانی آن به صورت گسترده ای در حال توسعه است و بدین ترتیب در آینده نزدیک به عنوان یک استاندارد مناسب و اقتصادی مطرح خواهد بود.
۵. ایکس ام ال داده ها را از اچ تی ام ال جدا می سازد.
۶. ایکس ام ال برای مبادله داده ها استفاده می شود.
۷. به کمک ایکس ام ال می توان به مبادله اطلاعات مالی در اینترنت پرداخت.
۸. ایکس ام ال تغییرات محیط را آسانتر می کند.
۹. ایکس ام ال داده ها را دسترس پذیرتر می سازد.
۱۰. ایکس ام ال برای ایجاد زبان جدید می تواند استفاده شود.

عناصر ایکس ام ال

- عناصر "ایکس.ام.ال." گسترش پذیر و مرتبط هستند. این عناصر قواعد نام گذاری ساده ای دارند.
- اسناد "ایکس.ام.ال." برای انتقال اطلاعات بیشتر می توانند گسترش داده شوند. عناصر "ایکس.ام.ال." بصورت والدین و فرزندان با هم مرتبط هستند. برای درک "ایکس.ام.ال." باید بدانید چگونه ارتباطات مابین عناصر "ایکس.ام.ال." نام گذاری و چگونه مندرجات عناصر توصیف می شوند.
- عناصر باید برچسب پایانی داشته باشند و به طور مناسبی جاگذاری شوند (خلیلی، ۱۳۸۴).

برخی از توانمندیهای ایکس.ام.ال

- ایکس.ام.ال. برای مواردی به کار می رود که اطلاعات مورد نظر به صورت ساخت یافته ای در دسترس هستند. اطلاعات ساخت یافته می توانند اطلاعاتی مانند متون و عکسها را با هدف و معنای آن اطلاعات تلفیق کنند. مثلاً در شرح یک کتاب می توان عنوان را به عنوان یک بخش اطلاعاتی عنوان کرد، روشن است که این نام در خارج از متن کتاب فاقد معنا و مفهوم است. این نام می تواند عنوان کتاب، شخصیت اول، بخشی از بیوگرافی یا نام نویسنده باشد، که فقط با بررسی ساختار نوشتار کتاب می توان معنای دقیق آن را تعیین کرد (خلیلی، ۱۳۸۴).
- ایکس.ام.ال. یک زبان نشانه گذاری بوده و در حوزه زبانهای مشابه اچ.تی.ام.ال. قرار می گیرد. این زبان همانند اچ.تی.ام.ال. از دو توضیح کوتاه^۵ استفاده می کند که همانند جلد یک کتاب در دو طرف بخشهای متن قرار می گیرند (شکل ۱). این توضیحات کوتاه اطلاعاتی در مورد متنی که در بر می گیرند ارائه می کنند.



- به طور کلی دو تفاوت اصلی بین اچ.تی.ام.ال. و ایکس.ام.ال. وجود دارد (Evjen 2008,260):
- اچ.تی.ام.ال. بر ظاهر متن و سایر اطلاعاتی که در یک جستجوگر نمایش داده می شوند تمرکز نمی کند، اما ایکس.ام.ال. به هدف یا معنای اطلاعات توجه دارد.

- اچ.تی.ام.ال. به یک مجموعه از پیش تعیین شده از توضیحات کوتاه محدود می شود ، اما کاربر می تواند با استفاده از ایکس.ام.ال. این توضیحات را خود ایجاد نماید
- در اینجا به بررسی یک مثال از زبان ایکس.ام.ال. می پردازیم . در این ایکس.ام.ال. شرحی از رمان تام جونز^۸، نوشته هنری فیلدینگ^۹ ذکر شده است . در این مثال و مثالهای دیگری که از زبان ایکس.ام.ال. ذکر شده اند، شماره های خطوط نیز جهت سادگی بررسی کدها در قسمت سمت چپ آمده اند ، که در واقع در ایکس.ام.ال. واقعی ذکر نمی شود.

```

01 <book>
02 <full-title>The history of Tom Jones A Foundling</full-title>
03 <author>
04 <name>Fielding Henry</name>
05 <birthplace>Glastonbury England </birthplace>
06 <birthyear>1707</birthyear>
07 </author>
08 <firstpublicationdate>1749 </firstpublicationdate>
09 <characters>
10 <characters>
11 <name>Tom Jones</name> <description>Hero of the story
12 </description> </character>
13
14 <character> <name> Sophia western</name>
15 <description>Heros true Love
16 </description>
17 <character> </characters>
18 <plot-summary>An orphan boy experiences
19 many adventures and Love affairs while growing up in eighteen-century England.</plot-summary>
20 <book>

```

در این مثال چندین ویژگی کلیدی از ایکس.ام.ال. آمده است:

- در ابتدای هر بخش اطلاعاتی یک توضیح کوتاه به شکل <tagname> و در انتهای آن نیز یک توضیح کوتاه به شکل </tagname> آمده است .
- اسامی توضیحات کوتاه معنا دار بوده و نوع اطلاعاتی که توضیح کوتاه شامل می شود را نشان می دهد .
- می توان توضیحات کوتاه را درون توضیحات کوتاه دیگر مورد استفاده قرار داد تا جزئیات بیشتری فراهم بیاید . مثلاً در خطهای ۰۳-۰۷ نام نویسنده ، محل تولد و سال تولد در توضیحات کوتاه اصلی </author> و <author> قرار دارند . یک زوج کوتاه که سایر بخشهای با توضیح را در بر می گیرند ، محفظه^{۱۰} نامیده می شوند . می توان مواردی را که دارای نوع یکسانی هستند در یک محفظه قرار داد. در خطوط ۰۹-۱۵ ، دو <character> درون محفظه <characters> قرار گرفته اند.
- در یک متن می توان از نامهای یکسانی با معنای متفاوت استفاده کرد . در این مثال توضیح کوتاه <name> برای نشان دادن نام نویسنده (خط ۰۴) و نامهای شخصیت‌های داستان (۱۱ و ۱۵) به کار رفته است. در اینجا هیچ گونه ابهامی وجود ندارد چرا که این نامها در محفظه های متفاوتی به کار رفته اند.

نتیجه گیری

ایکس.ام.ال. یک استاندارد صنعتی قدرتمند و انعطاف پذیر برای انتقال اطلاعات ساخت یافته است. این زبان توسط کنسرسیوم شبکه جهانی به عنوان ابزاری برای به اشتراک گذاردن اطلاعات از طریق اینترنت ارائه شده است. مجموعه مشخصات ایکس ام ال توسط یک گروه کاری در کنسرسیوم وب جهانی منتشر شده است و توسط تعدادی از شرکت های پیشرو در صنعت کامپیوتر پشتیبانی می شود.

می توان برای شرح تمام کتابهایی که توسط یک نویسنده نوشته شده اند، از یک ساختار ایکس.ام.ال. یکسان استفاده کرد، به راحتی این ساختار را تغییر داده تا کتابهایی را که توسط چندین نویسنده نوشته شده اند را در بر بگیرد . موضوع دیگری که باید مورد توجه قرار داد این حقیقت است که بجز فرمت و محل قرار گرفتن توضیحات کوتاه، ساختار کلی قرار گرفتن آنها در یکدیگر در تمام موارد دیگر فرضی است. البته می توان اسامی متفاوتی را برای این کوتاه مورد استفاده قرار داد، اطلاعات را به شکل دیگری دسته بندی نمود یا اطلاعات متفاوتی در کنار هم قرار داد. روشن است که خواندن و درک مثال فوق برای انسانها به سادگی امکان پذیر است . در صورتی که بخواهیم مثال فوق را در یک کاتالوگ پیوسته به کار ببریم، باید در مورد اسامی و ساختار توضیحات کوتاه و نوع اطلاعاتی که مورد توجه هستند ، توافق کنیم . در حقیقت استانداردها در چنین مواردی مطرح می شوند .

منابع

خلیلی، لیلیا ۱۳۸۴. مقدمه ای بر ایکس.ام.ال. (XML). مجله الکترونیک نما: ۴(۴).

http://www.irandoc.ac.ir/data/e_j/vol4/khalili.htm

والین، دن ۱۳۸۵. آموزش کاربردی ایکس.ام.ال. برای برنامه نویسان ASP.NET. ترجمه مهرداد توانا، سعید هراتیان، تهران: گروه مهندسی - پژوهشی ساحر.

Berners-Lee, Tim; Fischetti, Mark. 2005 Weaving the Web: the past, present and Future of the World Wide Web by its inventor. London: Orion Business,

Evjen, Bill 2008. XML. web services for ASP.NET, Newyork: Wiley.

Fitzgerald, Michael 2007, Building B2B Applications With XML: A Resource Guide, New York: John Wiley and sons.

Kurt Cagle 2003 , Begining XML.Indianapolis: Worx.

¹ Extensible Markup Language

² World Wide Web Consortium

³ Tag

⁴ Semantic Web

⁵ Advanced Distributed Learning (ADL)

⁶ AICC(Aviation Industry CBT Committee)

⁷ Tag

⁸ Tom jones

⁹ Henry Fielding

¹⁰ containers