

О.Г. Хамула

XML –ПРОГРАМУВАННЯ ЕЛЕКТРОННИХ ВИДАНЬ

Формування інформаційного суспільства спирається на сучасні інформаційні і телекомунікаційні технології та на технології зв'язку. Інформаційне суспільство ґрунтується на єдиному інформаційно-телекомунікаційному просторі, побудова якого вимагає широкого використання цифрових технологій для всіх видів інформаційних ресурсів. Цифрові технології застосовуються передусім у процесах створення інформаційних ресурсів. Використання настільних видавничих систем докорінно змінило всю видавничу галузь. З'явився новий тип інформаційних ресурсів масового призначення – електронні (мультимедійні) видання. Завдяки World Wide Web Інтернет перетворився у найпотужніший канал розповсюдження цих видань. Популярність WWW і невід'ємної її частини – HTML – безумовно, стала причиною підвищеної уваги до систем гіпертекстової розмітки документів – електронних (мультимедійних) видань. Хоча поняття гіпертексту було введене В. Бушем ще у 1945 р., а в 60-х роках стали з'являтися перші додатки, усе-таки пік зацікавленості цією технологією припадає на той період, коли виникла реальна потреба в механізмі об'єднання множини інформаційних ресурсів, у забезпеченні можливості створення і перегляду нелінійного тексту. І прикладом реалізації даного механізму послужило павутиння WWW.

Загалом, мова розмітки документів – це набір спеціальних інструкцій, так званих дескрипторів, або команд, призначених для формування в документах певної структури і визначення відношень між різноманітними елементами цієї структури. І хоча HTML у порівнянні з іншими коли-небудь запропонованими мовами електронних публікацій має певні переваги, проте занадто убогий: по суті, він говорить броузеру (програмі-оглядачу), як розмістити на сторінці текст, картинки, кнопки, що, у свою чергу, ускладнює розробку Web-сайтів. Назріли й інші проблеми в мережі. По-перше, Інтернет як супершвидкісна мережа часто-густо поводить себе гірше черепахи, а по-друге, незважаючи на те, що в ній присутня мало не вся інформація, знайти тут щось потрібне нерідко буває досить важко.

Розв'язання цих проблем покладається на створену і зареєстровану міжнародною організацією W3C у 1998 р. мову розмітки документів XML (Xtensible Markup Language). За декілька років офіційного існування вона привернула до себе увагу з боку як розробників, так і користувачів Інтернету. Сьогодні уже з'являються нові мови, створені на основі XML, виникають численні Web-сервери, які використовують цю технологію для організації інформації, що зберігається на них.

XML, як і HTML, – це система дескрипторів (елементів), що описують компоненти документа. Вона – не продовження HTML чи мова вищого рангу, а одна з підмножини того, що називається стандартною узагальненою мовою розмітки SGML.

Процес створення XML-документа дуже простий, вимагає лише базових знань HTML і розуміння тих завдань, які хочемо виконати, використовуючи XML як мову розмітки. При створенні XML-документів беруться до уваги такі вимоги:

- у заголовку документа розміщується оголошення XML із зазначенням мови розмітки документа, номера його версії та додаткової інформації;

- кожен відкриваючий дескриптор, що визначає певну область даних у документі, обов'язково повинен мати свого „напарника, що закриває”, тобто, на відміну від HTML, не можна опускати дескриптори, які закривають;

- враховується реєстр символів;

- значення атрибутів, використовуваних при визначенні дескрипторів, повинні розміщатися в лапках;

- інформація, що знаходиться між початковим і кінцевими дескрипторами, розглядається як дані, і тому враховуються всі символи форматування (прогалени, переходи рядків, табуляції не ігноруються, як у HTML).

Використання мови XML має свої переваги. У розробників з'являється унікальна можливість визначати власні команди, що забезпечує більш ефективно одержання даних, які розміщуються в документі. Автор документа створює його структуру, будує необхідні зв'язки між елементами, використовуючи команди, що задовольняють його вимоги, і домагається такого типу розмітки, яка потрібна для здійснення пошуку, перегляду й аналізу документа.

XML може використовуватись як універсальна мова запитів до сховищ інформації. А стандарт XML-QL (або XQL) уже

в недалекому майбутньому складе серйозну конкуренцію SQL. XML-документи можуть виступати також як унікальний інструмент збереження даних, що містить у собі одночасно засоби для опрацювання інформації та подання її клієнтові. У цій області одним із перспективних напрямків є інтеграція Java- і XML-технологій.

XML дозволяє здійснювати контроль за коректністю даних, що зберігаються в документах, перевіряти ієрархічні співвідношення у середині документа і встановлювати єдиний стандарт на структуру документів, вмістом яких можуть бути найрізноманітніші дані. Це означає, що його можна використовувати при побудові складних інформаційних систем, у котрих дуже важливим є обмін інформацією між різними додатками, які працюють в одній системі. Створюючи структуру механізму обміну інформацією з самого початку роботи над проектом, менеджер може застрахувати себе в майбутньому від багатьох проблем, пов'язаних із несумісністю використовуваних різноманітними компонентами системи форматів даних.

Не складні і програми-редактори XML-документів. Крім того, вільно поширюються всілякі програмні продукти, призначені для роботи з XML-документами. Сьогодні XML підтримується програмами Microsoft Internet Explorer 4.0, у новій версії Netscape Communicator і в бета-версіях IE5. Читач може навіть і не здогадатися, що має справу з XML, а не з HTML, якщо тільки не зверне увагу, що сайти з XML швидші і простіші у користуванні.

Але є і зворотний бік медалі. Йдеться про те, що мова розмітки XML дозволяє нам створювати індивідуальні дескриптори, документуючи їх за допомогою визначень типів документів (Document Type Definition, DTD) або схеми XML, а після того обмінюватися з іншими джерелами. І може виявитися, що інші використовують ті ж самі імена для елементів і атрибутів, але опираються на інакші DTD. Отже, створюючи нову мову розмітки на XML, потрібно домовитися, які будуть дескриптори, як вони можуть поєднуватися один з одним і яким чином повинні опрацьовуватись. Одним із недоліків є також те, що старі версії програм оглядачів не в змозі переглядати документи, написані мовою XML.

Таким чином XML, як і будь-яка нова технологія, потребує освоєння і доопрацювання. Ця мова, на наш погляд, прийшла

надовго і повинна полегшити діяльність розробників електронних видань.

1. Бублик В. Від електронних журналів до гібридних медій // Палітра друку. 2000. №4.
2. <http://www.vbxml.com/xml>.
3. http://www.xml_101.com/.
4. <http://www.nsu.ru/resource.html>.