1. Introduction to XML

XML 개요

□ Markup 언어

- markup을 사용하여 문서의 내용을 구조적으로 구분함으로써 내용을 쉽게 이해할 수 있도록 함
- markup 언어는 문서를 조직화하고 구조화시켜 문서의 내용을 정확하게 표현 및 전달할 수 있음

XML 개요

XML (eXtensible Markup Language)

- A de facto standard language for data representation and interchange over the Internet
- eXtensible: 확장가능 → 기존에 없던 것을 새로 만들 수 있음
- Markup Language: 문서의 내용(content)에 부가적인 정보를 표시할 수 있는 언어
 - » 예: HTML
 - ✓ stylistic markup: , , <i>, , ...
 - ✓ structural markup: , <a>, <div>, ...
 - ✓ semantic markup: <title>, <body>, ...

■ XML → Meta markup language

- 새로운 markup 언어를 정의할 수 있는 언어
 - » 의미를 갖는 엘리먼트 및 속성들의 집합과 사용 규칙을 정의

XML 개요

Generalized Markup

XML markup은 문서의 display format(출력 양식) 정보가 아닌,
 문서의 구조와 의미를 나타냄



Song>
<title>California love</title>
<artist>Tupac</artist>
<publisher>R record</publisher>
<length>4:30</length>
<date>2011/10/10</date>
</song>

XML 문서

XML 개요

□ 확장가능성

HTML 문서의 예 <html> <body> < XML 기조</td> >홍길동 </to> </html>

사용자는 기 정의된 태그 집합만 사용할 수 있음

XML 문서의 예 <?xml version="1.0"?> <booklist> <book> <title>XML 기초</title> <author>홍길동</author> </book> </booklist>

사용자가 새로운 언어(태그 집합, 문법)을 정의하고 그것을 이용하여 문서를 작성할 수 있음

5

SGML, XML, HTML, XHTML

SGML(Standard Generated Markup Language)

- 새로운 마크업 언어를 정의하기 위한 대표적인 Meta Markup Language
- 1986년 ISO에서 국제표준으로 채택됨
- 문법이 복잡하고 응용 S/W를 구현하기 어려워 공공기관 및 출판업계에서 제한적으로 사용됨

XML

- SGML의 부분집합으로, SGML의 장점을 취하고 복잡성을 배제하여 개발된 Meta Markup Language
- 1996년 W3C에서 개발 (http://www.w3.org/XML)
- 일반적으로 "XML 문서"라는 말은 XML로 정의된 마크업 언어로 작성된 문서를 뜻함

HTML

- SGML로 정의된 웹 문서 작성용 마크업 언어

XHTML

- XML로 정의된 웹 문서 작성용 마크업 언어

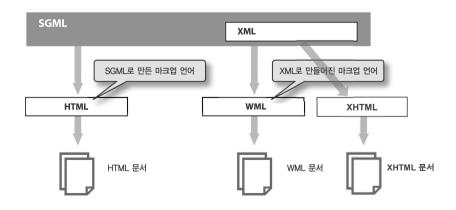
XML 개요

Meta markup language

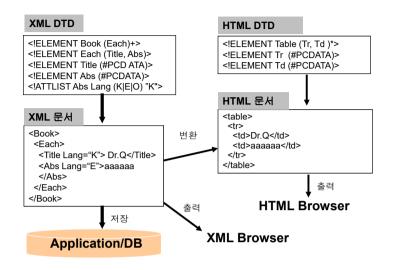
- DTD. XML Schema 등의 기술을 이용하여 새로운 마크업 언어를 정의 가능

```
<schema
targetNamespace="http://www.example.com/Memo"
xmlns:m ="http://www.example.com/Memo"
xmlns="http://www.w3.org/2001/XMLSchema">
                                                   <?xml version="1.0"?>
<element name="memo" type="m:memoType"/>
                                                   <memo priority="important">
                                                    <to>John</to>
<complexType name="memoType">
                                                    <subject>CS1111</subject>
 <sequence>
                                                    <message>
  <element name="to" type=string/>
                                                       Don't forget to attend the class
  <element name="subject" type=string/>
                                                       <emphasis>on Friday</emphasis>
  <element name="message" type="m:msgType"/>
                                                       Good luck to you.
  <element name="from" type=" string"/>
                                                    </message>
 </sequence>
                                                    <from>Tomas</from>
 <attribute name="priority" type="string"/>
                                                   </memo>
</complexType>
</schema>
```

SGML, XML, HTML, XHTML



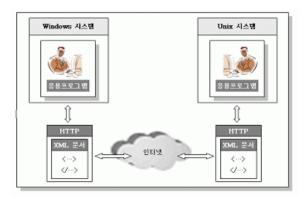
XML vs. HTML



XML 사용 분야

Electronic Data Interchange

 이기종 시스템 간의 정보 공유를 위한 데이터 교환 프로토콜로 XML 이용



XML 사용 분야

- □ 인터넷 상에서 **데이터 표현 및 교환**을 위한 산업계 표준(de facto standard) **언어/문서 양식**으로 사용됨
 - Electronic Data Interchange (EDI)
 - » 시스템 간 데이터 교환을 위한 공통 포맷으로 사용
 - Business to Business Integration
 - » 기업간 비즈니스 애플리케이션의 통합 (EAI)
 - » XML Web Services
 - Advanced Information Management System
 - » 모든 유형의 데이터 통합 관리
 - » Co-Work
 - » 지식관리시스템
 - Document Publishing Applications
 - » E-Book (EPUB)
 - Advanced Search System
 - » 구조 검색, 상품 카탈로그 검색

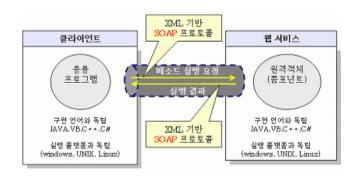


10

12

XML 사용 분야

- □ Business to Business(B2B) Integration
 - Web Services 기술을 이용한 분산 애플리케이션의 연동 및 통합
 - » Simple Object Access Protocol(SOAP)을 이용한 원격 객체 호출



History of XML

- □ ISO, 1986
 - SGML & SGML DTD
- W3C
 - XML 1.0 & XML DTD (Recommendation), 1998/02
 - XML 1.0 & XML DTD (Second Edition), 2000
 - XML 1.0 (Third Edition), 2004/02
 - XML 1.0 (Fourth Edition), 2006/08
 - XML 1.0 (Fifth Edition), 2008/11
 - » http://www.w3.org/TR/2008/REC-xml-20081126/
 - XML 1.1 (Recommendation), 2004/02
 - XML 1.1 (Second Edition), 2006/08
 - Namespaces in XML 1.0 (Third Edition), 2009/12
 - XML Schema Definition Language (XSD) 1.1 Part 1: Structures, 2012/04
 - XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes, 2012/04

13

DTD & XML Schema

- XML 문서의 구조 정의

```
<students>
                          student.xml
                                                                           student.dtd
                                         <!ELEMENT students (student*)>
 <student>
   <sno>s100</sno>
                                         <!ELEMENT student
                                                   (sno,name,age,phone,address)>
   <name>홍길동</name>
                                         <!ELEMENT sno (#PCDATA)>
   <age>26</age>
   <phone>02-123-8989</phone>
                                         <!ELEMENT name (#PCDATA)>
                                         <!ELEMENT age (#PCDATA)>
   <address>서울시 월곡동</address>
                                         <!ELEMENT phone (#PCDATA)>
 </student>
</students>
                                         <!ELEMENT address (#PCDATA)>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
                                                                           student.xsd
<xsd:element name="student">
 <xsd:complexType>
  <xsd:sequence>
   <xsd:element name="sno" type="xsd:string"/>
   <xsd:element name="name" type="xsd:string"/>
   <xsd:element name="age" type="xsd:string"/>
   <xsd:element name="phone" type="xsd:string"/>
   <xsd:element name="address" type="xsd:string"/>
   </xsd:sequence>
 </xsd:complexType>
</xsd:element>
</xsd:schema>
```

XML 관련 기술

XML is a family of technologies

- Schema: DTD, XML Schema

Rendering: XSL, CSS

Transformation: XPath, XSLT, XSL-FO

Query: XPath, XQuery

Publishing: XSL-FO

Processing: XML Pipeline(XProc)

Security: XML Encryption, XML Signature, XKMS

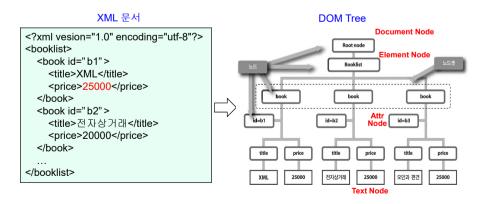
Components: XLink, XPointer, XInclude, XForms, etc.

Internationalization: ITS

Parsing API: DOM, SAX

XPath

- XML 문서의 엘리먼트나 속성을 접근하기 위한 경로(path) 표현 방법

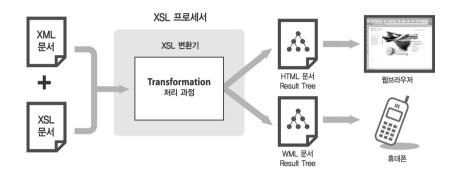


- XPath 식의 예와 실행 결과
 - ▶ /booklist/book/title
 → XML, 전자상거래, 오만과 편견
 - > //book[title='XML']/price → 25000

^{*} XML Technology (http://www.w3.org/standards/xml/) 참조

eXtensible Stylesheet Language – Transformation (XSLT)

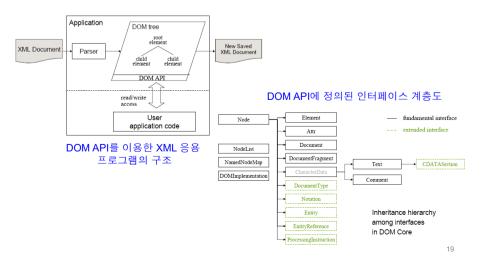
- XML 문서의 변환 규칙 및 출력 양식 지정



17

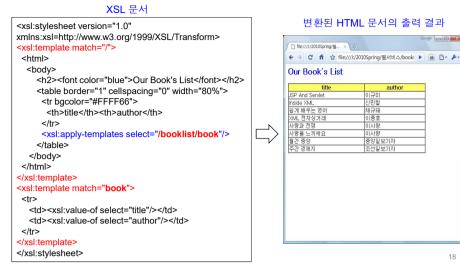
Document Object Model (DOM)

- XML 문서를 해석(parsing) 및 이용하기 위한 tree-based API



eXtensible Stylesheet Language – Transformation (XSLT)

- XML 문서의 변환 규칙 및 출력 양식 지정



Document Object Model (DOM)

- DOM API를 이용한 XML 프로그램 작성 예
 - » XML 문서에 포함된 모든 book 엘리먼트에 publishDate 자식 엘리먼트와 날짜 데이터 추가

```
public static void addPublishDate(Node n)
                                                        // n은 root node부터 시작
   if (n.getNodeType() == Node.ELEMENT NODE) {
                                                        // 엘리먼트를 처리함
       if (n.getNodeName().equals("book")) {
                                                        // book 엘리먼트 발견
           Document doc = n.getOwnerDocument();
                                                        // 문서 객체를 찾음
           Element date = doc.createElement("publishDate"); // 엘리먼트 노드 생성
           Text txt = doc.createTextNode("20130501");
                                                         // text node 생성
           date.appendChild(txt);
                                       // <publishDate>20130501</publishDate>
                                      // 현재 엘리먼트의 마지막 자식 엘리먼트로 추가
           n.appendChild(date);
       for (Node ch = n.getFirstChild(); ch != null; ch = ch.getNextSibling()) {
                                           // 각 자식 엘리먼트에 대해 recursion 수행
           addPublishDate(ch);
                                                                              20
```

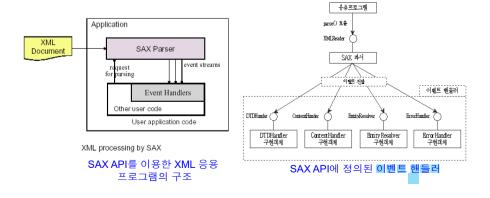
HTML DOM

- HTML 문서의 내용을 DOM tree로 표현하고 HTML DOM API를 이용해서 이용 및 조작
- 활용 예



Simple API for XML (SAX)

- XML 문서를 해석(parsing) 및 이용하기 위한 event-based API

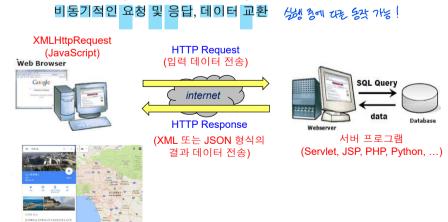


HTML DOM

```
function addOnClickHandlers() {
 var cdsDiv = document.getElementById("cds");
 var cdImages = cdsDiv.getElementsByTagName("img");
 for (var i=0; i<cdImages.length; i++) // CD 리스트의 각 <img> 마다
   cdImages[i].onclick = addToTop5;
                                       // onclick event handler 추가
function addToTop5() {
                                      // 각 <img> 클릭 시 호출됨
 var imgElement = this;
                                      // 이 함수를 호출한 <img> 엘리먼트
 var top5Element = document.getElementById("top5");
top5Element.appendChild(imgElement); // <img>가 <top5>의 자식으로 변경됨
 imgElement.onclick = null;
                                      // onclick event handler 제거
function startOver() {
                                      // Start Over 버튼 클릭 시 호출됨
 var top5Element = document.getElementById("top5");
 var cdsElement = document.getElementById("cds");
 while (top5Element.hasChildNodes()) {
   var firstChild = top5Element.firstChild; // <top5> 밑의 각 <img>를
   cdsElement.appendChild(firstChild);
                                          // <cds>의 자식으로 변경
 addOnClickHandlers(); }
```

Asynchronous JavaScript & XML (Ajax)

- 웹 브라우저(클라이언트 프로그램)과 서버 프로그램 간의



Ajax 활용 예

- 입력 창에 글자를 입력할 때마다 서버 프로그램에 비동기적으로 추천 키워드들을 요청
- 서버 프로그램은 입력된 단어에 대한 추천 키워드들을 검색 및 반환

Start typing a name in the input field below:

First name: el x	
Suggestions: Eva, Eve, Evita, Elizabeth, Ellen	
Ţ	
Start typing a name in the input field below:	Start typing a name in the input field below:
First name: ell x	First name: elis x
Suggestions: Elizabeth, Ellen	Suggestions: no suggestion

21

Ajax 활용 예

- 서버 프로그램 (JSP)

```
<%
  // 이름들의 배열 정의 (→ 또는 데이터베이스 검색)
  String[] names = { "Anna", "Brittany", "Cinderella", "Diana", "Eva", "Fiona", "Gunda", ... };
  String q = request.getParameter("q");
                                          // URL에서부터 q 파라미터 값 (입력 값) 추출
  String hint = null;
  // g가 빈문자열이 아니면 배열에서 모든 이름 값 조회
  if (q != null) {
      q = q.toLowerCase();
                                                                // 소문자로 변환
      for (int i = 0; i < names.length; i++) {
                                                                // 각 name 조사
                                                          // i번째 이름이 q로 시작하면
         if (names[i].toLowerCase().startsWith(q) == true) {
             if (hint == null) { hint = names[i]; }
                                                                // hint 문자열에 저장
             else { hint = hint + ", " + names[i]; }
                                                 // 응답 메시지로 결과 출력 (text string)
  if (hint != null) out.print(hint);
  else out.print("no suggestion");
%>
```

Ajax 활용 예

- HTML & JavaScript 프로그램

```
<html><head><script>
function showHint(str) {
                                                     // str: 입력 필드의 값
                                                     // 입력 값이 없으면
if (str.length == 0) {
  document.getElementBvId("txtHint").innerHTML = "":
                                                     // txtHint의 내용을 clear
  return:
} else {
  var xhttp = new XMLHttpRequest():
                                                 // 서버 응답에 대한 callback 함수
  xhttp.onreadystatechange = function() {
     if (this.readyState == 4 && this.status == 200) // 응답 메시지를 txtHint 값으로 설정
       document.getElementBvld("txtHint").innerHTML = this.responseText:
  xmlhttp.open("GET", "gethint.jsp?q=" + str, true); // 입력 값을 query string으로 전송
  xhttp.send():
                                                               키보드로 한 문자 입력 시
                                                               showHint() 함수를 호출하고
</script></head>
                                                               입력된 문자열을 전송
<body>
<b>Start typing a name in the input field below:</b>
<form> First name: <input type="text" onkeyup="showHint(this.value)" /></form>
Suggestions: <span id="txtHint"></span> ... // 키워드 hint 출력 장소
```

XML로 정의된 Markup 언어

Non-textual Data

<u> </u>	toxtuui Butu	
Language	Purpose	Document
SMIL	Multimedia documents	- <u>Synchronized Multimedia Integration</u> <u>Language (SMIL 3.0), R, Dec. 2008</u>
MathML	Mathematical notation	- A MathML for CSS profile, R, June 2011 - Mathematical Markup Language (MathML) Version 3.0, R, Oct. 2010
SVG X3D	Vector graphics Extensible 3D Graphics	- Scalable Vector Graphics (SVG) 1.1 (2nd Edition), R. Aug. 2011 - ISO/IEC 19775-1.2:2008: X3D Architecture and base components Edition 2
VoiceXML CCXML SSML	Voice markup and telephony call control; to enable access to the Web using spoken interaction	- Voice Extensible Markup Language (VoiceXML) 2.1, R, June 2007 - Voice Browser Call Control: CCXML Version 1.0, R, July 2011 - Speech Synthesis Markup Language (SSML) Version 1.1, R, Sept. 2010
EmotionML	To describe human emotions	- Emotion Markup Language (EmotionML) 1.0, CR, May 2012

XML로 정의된 Markup 언어

Web Publishing

	•	
Language	Purpose	Document
XHTML	Reformulation of HTML 4.0 i n XML	- XHTML™ 1.0: The Extensible HyperText Markup Language (Second Edition), A Reformulation of HTML 4 in XML 1.0, R, Jan. 2000, revised 1 August 2002
XML Events	To represent asynchronous occurrences, such as mouse clicks, in XHTML or in other XML markup	- XML Events, An Events Syntax for XML, R. Oct. 2003
XForms	Web forms	-XForms 1.0 (Third Edition), R, Oct. 2007
XHTML-Print	Simple XHTML suitable for printing from mobile devices as well as for display	- XHTML-Print - Second Edition, R, Nov. 2010
Polyglot Markup	For writing documents that serve bothas HTML 5 and XML documents	- Polyglot Markup: HTML-Compatible XHT ML Documents , TR, March 2012

29

XML로 정의된 Markup 언어

Web Communication and Services

Language	Purpose	Document
XML-Signature	Associating digital objects by digital signatures in XML format	- XML-Signature Syntax and Processing (Second Edition), R, June 2008
XML Encryption	Encrypting data and represen ting the result in XML	- XML Encryption Syntax and Processing, R, Dec. 2002 - Decryption Transform for XML Signature, R, Dec. 2002
XKMS	Protocol for distributing and registering public keys	-XML Key Management Specification (XKMS) Version 2.0, R, June 2005
SOAP	Protocol	- <u>SOAP Version 1.2 Part 0: Primer (Second Edition)</u> , R, April 2007
WSDL	To describe Web services	- Web Services Description Language (WSDL) Version 2.0 Part 0: Primer, R. June 2007
SML	Service modeling	- <u>Service Modeling Language, Version 1.1,</u> R. May 2009

XML로 정의된 Markup 언어

Metadata and Semantic Web

Language	Purpose	Document
RDF	Metadata for Web resources	- Resource Description Framework (RDF): <u>Concepts and Abstract Syntax, R, Feb.</u> <u>2004</u> - <u>RDFa Core 1.1, R, June 2012</u>
RDF Schema	To describe RDF vocabularies	- <u>RDF Vocabulary Description Language</u> 1.0: RDF Schema, R, Feb. 2004
OWL	Semantic markup language for publishing and sharing ontologies on the Web	- OWL 2 Web Ontology Language. Primer, R, Oct. 2009
SPARQL	Query language for RDF	- <u>SPARQL Query Language for RDF, R, Jan. 2008</u>
RIF	Rule Exchange Format	- RIF Core Dialect, R, June 2010

XML로 정의된 Markup 언어

□ OFX(Open Financial eXchange) 문서 예 🐉 업무

 Request of an open-ended statement download (no start/end dates) for savings account # 098-121

```
<OFX>
  <SIGNONMSGSRQV1>
  </SIGNONMSGSRQV1>
  <BANKMSGSRQV1>
    <STMTTRNRQ>
      <TRNUID>23382938
      <STMTRQ>
        <BANKACCTFROM>
          <BANKID>987654321
          <ACCTID>098-121
          <ACCTTYPE>SAVINGS
        </BANKACCTFROM>
        <INCTRAN>
          <INCLUDE>Y
        </INCTRAN>
      </STMTRQ>
    </STMTTRNRO>
  </BANKMSGSRQV1>
</OFX>
```

XML로 정의된 Markup 언어

□ MathML 문서 예

$$x = \frac{-b \pm \sqrt{b^2 - 4aa}}{2a}$$

```
<math xmlns="http://www.w3.org/1998/Math/MathML">
 <apply><in/>
   <ci>x</ci>
   ceet >
     <apply><divide/>
       <apply><plus/>
         <apply><minus/>
           <ci>b</ci>
          </apply>
          <apply><root/>
           <apply><minus/>
              <apply><power/>
               <ci>b</ci>
               <cn>2</cn>
              </apply>
              <apply><times/>
               <cn>4</cn>
                <ci>a</ci>
                <ci>c</ci>
              </apply> </apply> </apply> </apply>
     <apply><times/> <cn>2</cn> <ci>a</ci> </apply> </apply>
     ... </set> </apply> </math>
```

33

XML로 정의된 Markup 언어

X3D(Extensible 3D Graphics)

ISO standard XML-based file format for representing 3D computer graphics

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE X3D PUBLIC "ISO//Web3D//DTD X3D 3.1//EN" "http://www.web3d.org/specifications/x3d-
3.1.dtd"> <X3D profile='CADInterchange' version='3.1' xmlns:xsd='http://www.w3.org/2001/XMLSchema-
instance' xsd:noNamespaceSchemaLocation =' http://www.web3d.org/specifications/x3d-3.1.xsd '>
<head> <component level='2' name='CADGeometry'/>
 <meta name='title' content='CadTeapot.x3d'/>
 <meta name='description' content='Teapot model demonstrating proper hierarchy of CAD nodes:</p>
CADLayer, CADAssembly, CADPart, and then CADFace.'/>
<meta name='creator' content='Alan Hudson, Don Brutzman'/> ...
</head>
<Scene>
 <Background skyColor='0.9607843 1 0.9607843'/>
 <Viewpoint description='Hello CAD teapot' position='0 0 35'/>
 <CADLayer DEF='TopLayer' name='Single topmost CADLayer for this model'>
   <CADAssembly name='Assembled Teapot'>
     <CADPart name='Body'>
        <CADFace name='Teapot body IFS mesh'>
          <Shape containerField='shape'>
             <Appearance DEF='APP01'><Material diffuseColor='0.3451 0.5608 0.8824'/>
             <IndexedFaceSet creaseAngle='1.0' coordIndex='0 5 6 0 5 6 -1 6 1 0 -1 1 6 7 -1 7 2 1 -1 2 7 8</p>
-1 8 3 2 -1 3 8 9 -1 9 4 3 -1 5 10 11 -1 11 6 5 -1 6 11 12 -1 12 7 6 -1 7 12 13 -1 13 8 7 -1 8 13 14 -1 14 9 8 -1
10 15 16 -1 16 11 10 -1 11 16 17 -1 17 12 11 -1 12 17 18 -1 18 13 12 -1 13 18 ... /> ...
```

XML로 정의된 Markup 언어

RSS(Rich Site Summary/Really Simple Syndication)

 Web feed formats used to publish frequently updated works - such as blog entries, news headlines, etc. - in a standardized format

```
<rss version="2.0" xmlns:dc="http://purl.org/dc/elements/1.1/">
  <channel>
     <title>동아닷컴 : 동아일보 IT/의학 뉴스</title>
                                                       동아닷컴 : 동아일보 ITI대학 노스용 때도
다음을 사용하여 되도 구독 (Mensible 보기 지르므로)
디피드를 구독한 때 항상이 리더 사용
     <copyright>Copyright donga.com</copyright>
     k>https://www.donga.com/</link>
     <description>동아닷컴::RSS 서비스 | 동아일보 IT/의학
     <language>ko</language>
     <pubDate>Sun, 03 Mar 2024 18:47:38 +0900</pubDate</pre>
      <title><![CDATA[ "더 오래가고 충전 빨라졌다"... SK온
      k>https://www.donga.com/news/Economy/artic
                                                      마스크 오프레에 =개도하이의 비역기 용제 외바
      <pubDate>Sun, 03 Mar 2024 18:40:00 +0900</pubD</pre>
      <description><![CDATA[<imq
src='https://dimg.donga.com/i/150/150/90/wps/NEWS/IMA
61.2.jpg' align='left' vspace='10' hspace='10'>SK온이 에너지 밀도를 높이면서 충
전 시간을 단축한 전기차 배터리를 선보인다. 전기차 주행거리를 늘리면서 충전 성능을 향상
시킨 것으로 전기차 사용 편의를 전반적으로 개선했다는 평가다. ... 11>
      </description>
      <dc:category>IT/의학</dc:category>
     </item>
     <item> ... </item> <item> ... </item> ... </channel> </rss>
                                                                                34
```

기타 XML 사용 예

Word processing formats

- OASIS's Open Document Format (http://www.odfalliance.org/)
- Microsoft's Open Office XML (http://www.openxmlcommunity.org/)
- Apple's iWork

Graphics formats

- Scalable Vector Graphics(SVG) (http://www.w3.org/Graphics/SVG/)
- Extensible 3D Graphics (X3D) (http://www.web3d.org/)
 - » পা: http://www.web3d.org/x3d/content/examples/HelloWorld.x3d,
 http://www.web3d.org/x3d/content/examples/K3dResources.html#Examples

Web feed formats

- RSS (http://en.wikipedia.org/wiki/RSS)
- ATOM (http://en.wikipedia.org/wiki/Atom)
 - » 예: https://rss.donga.com/, https://news.google.com/rss/, http://section.blog.naver.com/

기타 XML 사용 예

- □ Electronic publications (e-book)
 - EPUB (http://idpf.org/epub/31)
- Messaging protocols/ data format
 - AJAX(Asynchronous JavaScript and XML)
 - XML-RPC, SOAP, XMPP, ...
 - RESTful Web Services, Open API
- Programming configurations
 - Deployment descriptor in Java EE (ex: web.xml)
 - Configuration in Struts/Spring framework
 - UI Layout definition in Android applications
 - Applications for Microsoft .NET Framework
- XML Data Repository at UW Database Group
 - http://www.cs.washington.edu/research/xmldatasets/

References

- Extensible Markup Language (XML) 1.0 (Fifth Edition)
 W3C Recommendation, 26 November 2008
 - http://www.w3.org/TR/REC-xml/
- W3C XML Technology
 - http://www.w3.org/standards/xml/
- XML (from Wikipedia)

37

- http://en.wikipedia.org/wiki/XML
- List of XML markup languages
 - http://en.wikipedia.org/wiki/List of XML markup languages
- XML Applications and Initiatives
 - http://xml.coverpages.org/xmlApplications.html