

제 1 장 XML 소개

1-1. XML의 어원적 뜻

1-2. 마크업 언어

1-3. 확장 가능의 의미

1-4. 메타언어로서의 XML

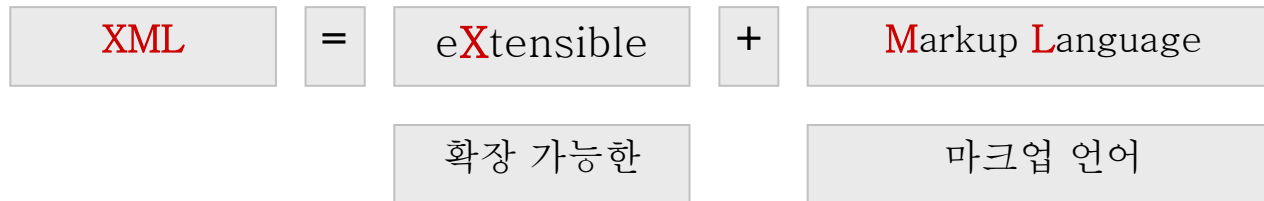
1-5. SGML과 XML 비교

1-6. W3C에서 하는 일

1-7. W3C의 XML 권고안 (Recommendation)

1-1. XML의 어원적 뜻

- XML(eXtensible Markup Language)은 “확장 가능한 마크업 언어”이다.



확장 가능

-기존에 없던 것을 새로 만들 수 있다.

마크업 언어

-문서의 내용을 구체적으로 전달하기 위한 추가적인 정보 표시 언어이다

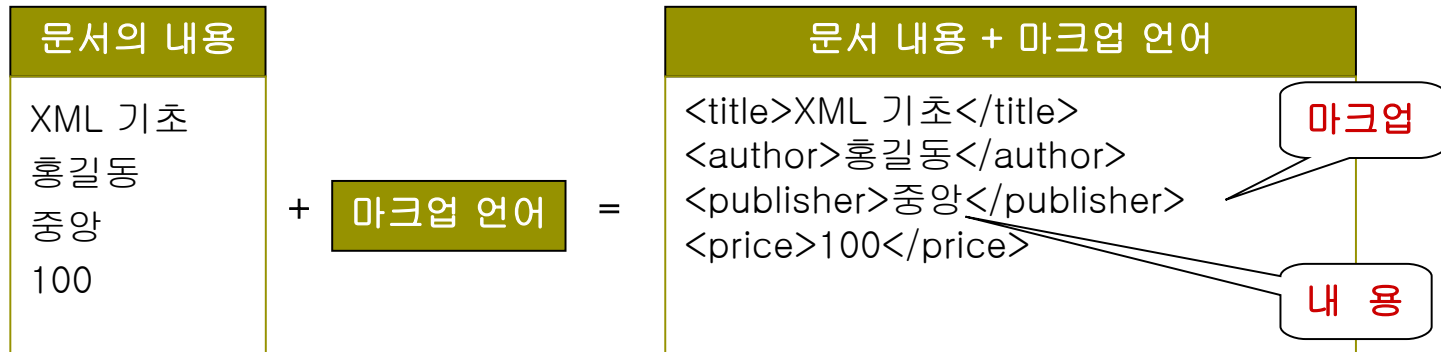
.

(예: HTML의 각종 태그)

XML은 “기존에 없던 것을 새로 만들 수 있는 마크업 언어”이다.
XML은 새로운 마크업 언어를 개발하기 위한 수단으로 사용된다.

1-2 마크업 언어

- **마크업 언어**란 문서의 내용에 추가적인 정보를 표시하기 위한 언어를 의미한다.
- 문서의 내용과 마크업 언어와의 관계



- 마크업을 사용하여 문서의 내용을 구조적으로 구분함으로써 내용을 쉽게 이해할 수 있게 된다.
- 마크업 언어는 문서의 내용을 조직화 하고 구조화 시켜 문서의 내용을 정확하게 전달한다.

1-3 확장 가능한 의미

- “확장 가능”이란 “기존에 없던 것을 새로 만들 수 있다.”는 의미이다.

고정 DTD

HTML 문서의 작성 예

```
<html>
  <body>
    <table>
      <tr>
        <td>XML 기초</td>
        <td>홍길동</td>
      </tr>
    </table>
  </body>
</html>
```

정해진 태그만 사용할 수 있음

사용자 정의
DTD, Schema

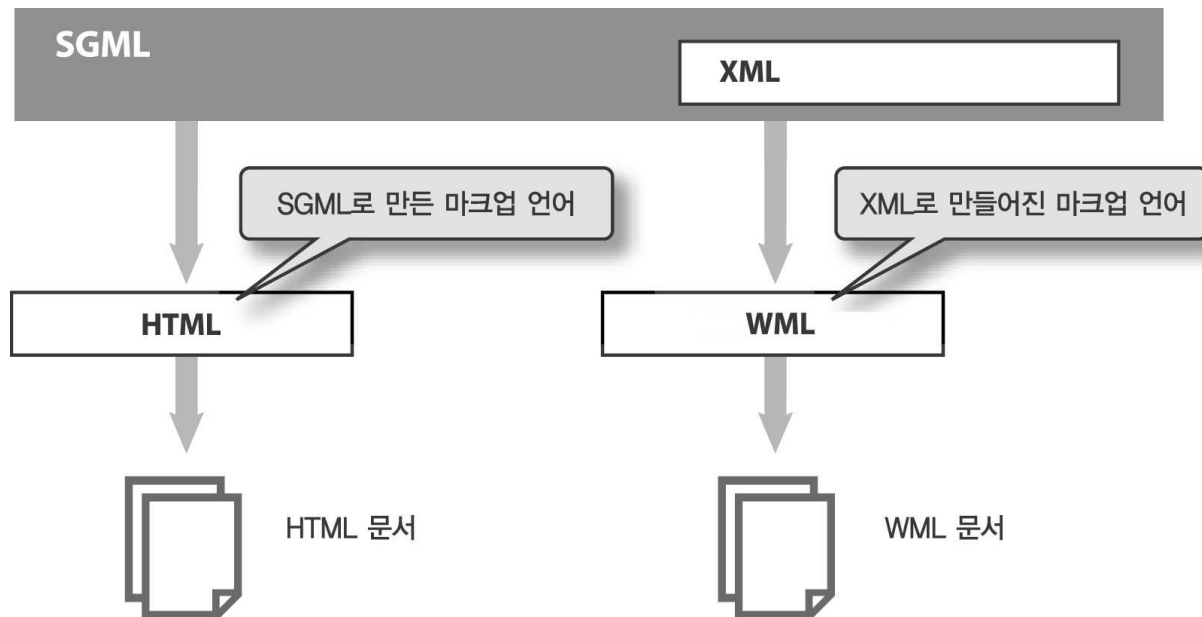
XML 문서의 작성 예

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

새로운 마크업을 개발하여 문서를
작성할 수 있음

1-4 XML은 메타언어

- XML은 **메타언어(Meta-Language)**로써 마크업 언어를 개발하기 위한 언어이다.
- 대표적인 메타언어는 **SGML(Standard Generated Markup Language)**이며, HTML은 SGML로 개발된 웹 문서 작성용 마크업 언어이다.
- XML은 SGML과 같은 메타언어이므로 새로운 마크업 언어를 만드는 도구로 사용된다. “XML로 문서를 작성한다”라는 말은 XML로 개발된 특정 마크업 언어로 문서를 작성하는 것을 의미한다.



1-5. SGML과 XML 비교

- **SGML**은 1986년 **ISO**(International **S**tandardization **O**rganization)에서 표준으로 발표.
- 문법이 복잡하고, 응용 S/W를 구현하기 어려워 주로 공공기관 및 출판업체에서 제한적으로 사용.
- 1996년 W3C의 **SGML Work Group**에서 SGML의 부분집합(Subset)으로서 SGML의 장점을 취하고, 복잡성을 배제한 XML을 개발
< W3C의 XML 개발 목표 >

XML은 인터넷에서 곧바로 사용 할 수 있어야 한다..

XML은 폭넓은 application을 지원해야 한다.

XML은 SGML과 호환성이 있어야 한다.

XML 문서들을 처리하는 프로그램을 쉽게 사용할 수 있어야 한다.

XML의 선택적 성능(feature)들이 최소화 하거나 가능하면 없어야 한다

.

XML 문서들은 사람이 이해하기 쉬워야 한다.

XML 디자인을 빨리 할 수 있어야 한다.

XML 디자인은 특정 양식에 따르고 함축적이어야 한다.

XML 문서들은 생성하기 쉬워야 한다.

XML 작성(markup)에서의 간결성(terseness)은 중요하지 않다.

1-6. W3C에서 하는 일

- 1994년 10월 Tim Berners-Lee(web의 창시자) 설립.
- W3C에서는 각종 스펙(Spec)을 연구하고 발표하는 일을 한다.

국제표준화기구 (ISO)에서 정한 규격 → 표준(standardization)
W3C에서 정한 규격 → 권고안(Recommendation)

- < W3C에서의 스펙 발표 진행 단계 >



1-7. W3C의 XML 권고안(Recommendation)

- 1997년 9월 노트(**Note**)로 시작
- 1997년 12월 후보 권고안(Propose Recommendation) 발표
- 1998년 2월 XML 1.0 권고안(**Recommendation**) 발표(한글화)
- 2004년 2월 XML 1.0 세 번째(**Third Edition**) 권고안 발표

<http://trio.co.kr/webrefer/xml/xml10.html>

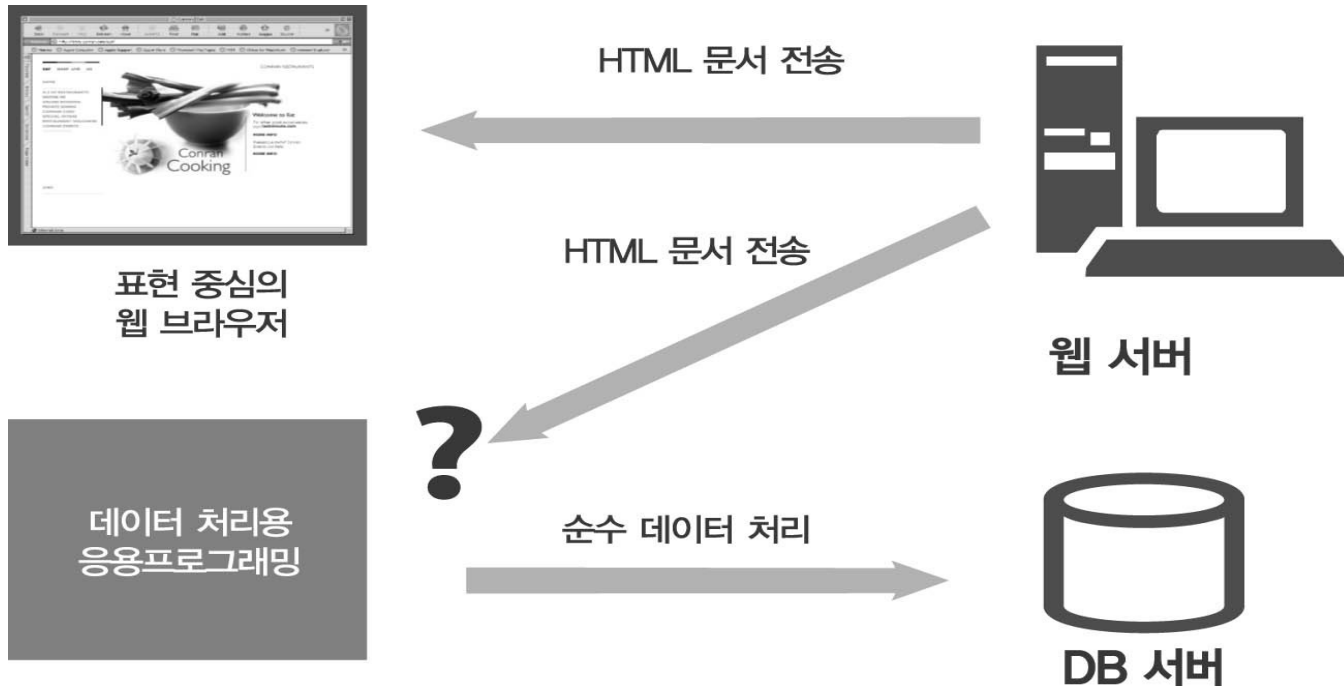
2. XML 문서 사용 용도

2-1. 데이터 전달용 문서

2-2. HTML에 대한 잘못된 오해

2-1. 데이터 전달용 문서

- 웹 서버가 데이터의 내용과 표현 형식을 HTML 문서에 함께 담아서 클라이언트로 보내면 웹 브라우저는 태그가 의미하는 표현 방법으로 그 데이터를 디스플레이 해준다.
- HTML 문서를 **표현 중심**이 아닌, **정보 전달**용으로 사용했을 때의 문제점
→ HTML 문서 내용에서 데이터를 추출할 수는 있지만, 데이터의 의미는 알기 어려움

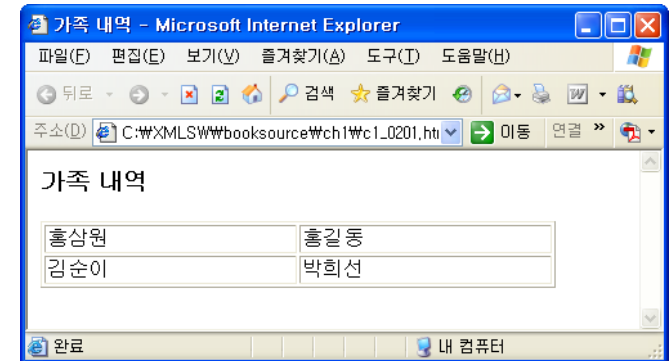


2-1. 데이터 전달용 문서(HTML의 한계점)

- HTML 문서는 순수 데이터를 추출하기 어렵다.

HTML 문서

```
<html>
  <head>
    <title>가족 내역</title>
  </head>
  <body>
    <h3>가족 내역</h3>
    <table border="1" width="400">
      <tr> <td>홍삼원</td> <td>홍길동</td> </tr>
      <tr> <td>김순이</td> <td>박희선</td> </tr>
    </table>
  </body>
</html>
```



- 데이터 추출에는 성공 했지만, 데이터가 의미하는 내용은 알 수 없다.
- HTML 문서를 데이터 처리용 응용프로그램에서 이용한다면 추출한 데이터를 적절히 이용할 수 없다.

2-1. 데이터 전달용 문서(계속)

< 해결 방법 >

- ① HTML에서 가족 관계를 나타내는 태그를 만든다. → 불가능(이미 태그들이 고정되어 있음)
- ② HTML 문서의 문제점을 극복한 XML 문서를 사용한다.

XML 문서

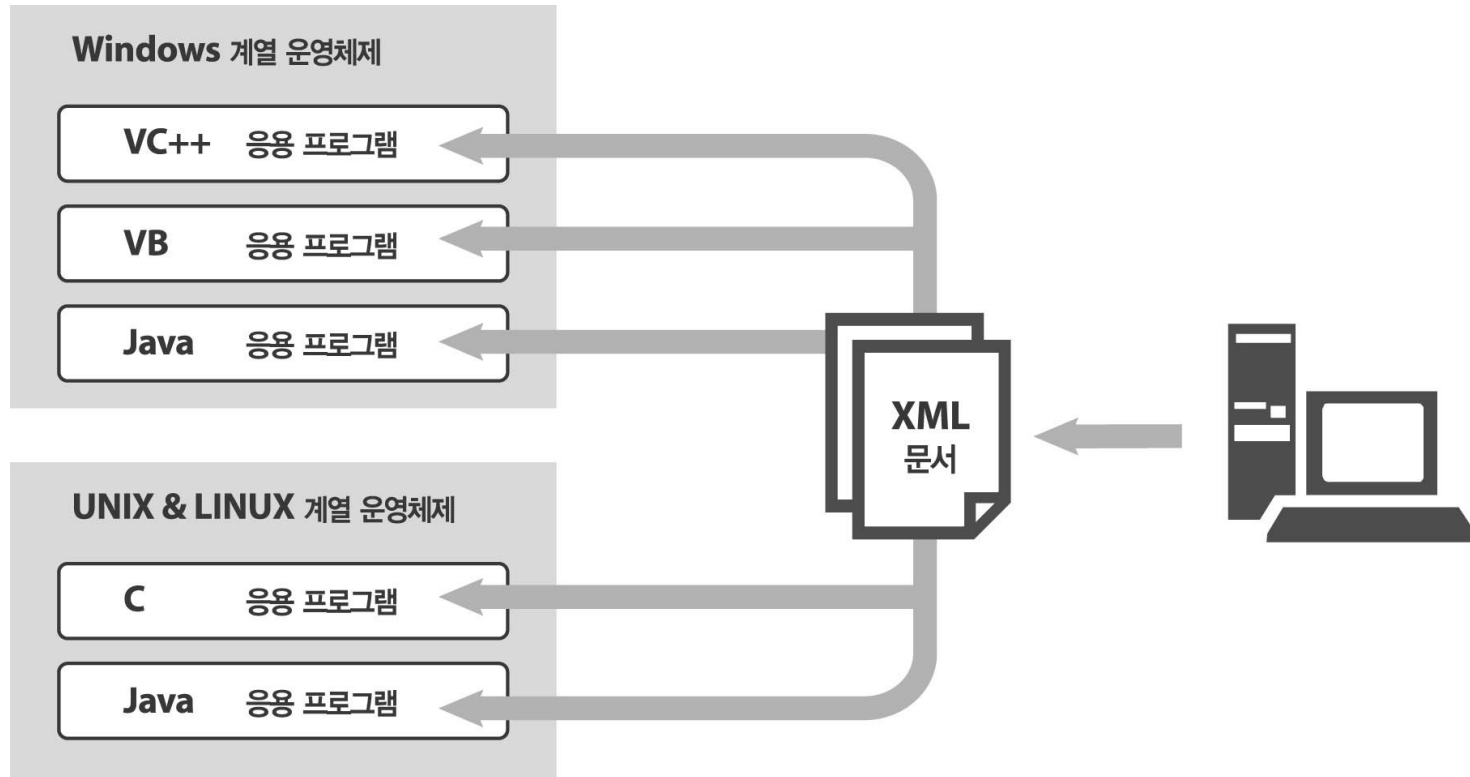
```
<?xml version="1.0"?>
<가족관계>
  <아버지>홍삼원</아버지>
  <어머니>김순이</어머니>
  <아들>홍길동</아들>
  <며느리>박희순</며느리>
</가족관계>
```

조직화, 구조화된
데이터 전달

→ XML 문서는 응용 프로그램으로 하여금 쉽고 정확하게 데이터를 처리할 수 있도록 데이터를 조직화되고 구조화 시킨다.

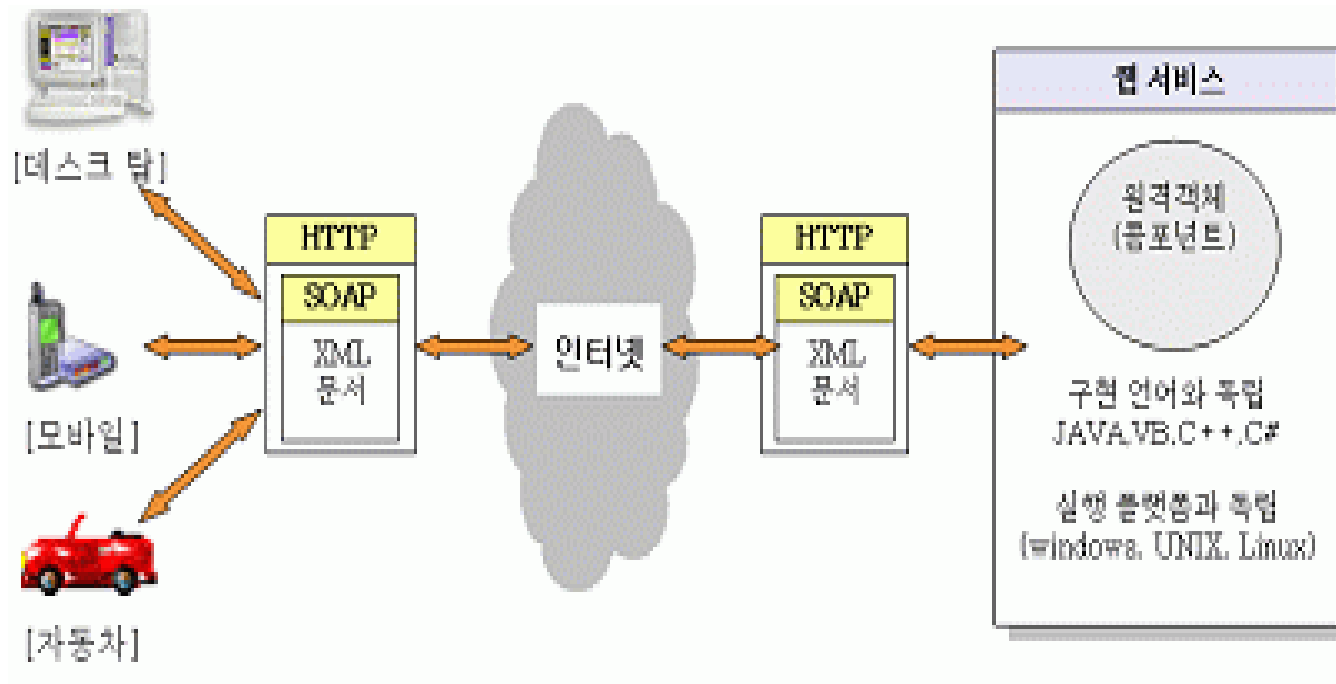
2-1. 데이터 전달용 문서(계속)

- 데이터를 처리하는 응용프로그램도 XML 문서가 Text형식의 **유니코드(Unicode)**로 작성되기 때문에 유니코드를 인식하는 모든 운영 체제상의 응용 프로그램과도 호환성 있는 문서가 된다.



2-2. XML 문서의 적용 분야

- XML 문서의 핵심은 시스템 통합
- 이기종 시스템 정보 교환 분야
- 웹 서비스 분야
 - 유비 쿼터스(Ubiquitous) : “언제 어디서나 동시에 존재한다”. Anywhere, Anytime, Any device → 3A



2-3. HTML에 대한 잘못된 오해

- XML이 HTML을 대신하지는 않는다. XML은 HTML의 대체 언어가 아니며, 이 두 가지 언어는 상호 보완적인 언어로 앞으로도 계속 사용될 것이다.
- XML은 HTML을 확장한 것이 아니다. XML이 데이터 전달용으로 사용될 경우 HTML 문서 보다 더욱더 좋은 해결책일 뿐이다.
- 전송용 데이터는 XML로 개발된 마크업 언어에 의해 XML 문서로 조직화, 구조화시켜 네트워크를 통해 원격 전송하면 좋다.
- 표현 중심의 문서를 작성할 때는 HTML이 가장 좋은 언어이다.