

실습 #1: DTD를 이용한 Markup 언어 설계

□ 주의사항

- 먼저 DTD에 대한 기본적인 이론을 잘 이해해야 함
- 강의노트 및 교재의 예제를 참고
- DTD를 작성하는 과정에서 XML 문서를 함께 작성 및 테스트하면서 점진적으로 완성하는 것이 바람직함
 - 유효성 검사 실시 및 오류 수정
- 한 가지 정답이 아니라 여러 가지 대안이 있을 수 있음
→ Design Choice

설계 방안

- R1: BookML 문서는 도서들의 목록과 도서 분류 체계로 구성된다.

```
<catalog>
  <bookList>
    ...
  </bookList>
  <bookCategory>
    ...
  </bookCategory>
</catalog>
```



```
<!-- root element 선언 -->
<!ELEMENT catalog (_____, _____)>
```

기본 요구사항

No	내용
R1	BookML 문서는 도서들의 목록과 도서 분류 체계로 구성된다.
R2	각 도서는 제목, 저자, 출판사, 출판일, 가격, 할인율, 페이지수, 평점, 분야(category)에 대한 정보를 갖고, 추가적으로 개정일, 표지그림, 책소개, 독자리뷰를 가질 수 있다. 번역서는 원제목과 역자 정보도 있다. 저자와 분야, 개정일은 여러 개가 존재할 수 있고, 독자리뷰는 최대 20개까지만 가능하다.
R3	도서의 식별자는 ISBN10 (isbn-xxxxxxxx-x) 또는 ISBN13 (isbn-978(또는 979)-xxxxxxxx-x) 형식을 사용하고, 도서마다 서로 다른 유일한 값을 갖는다.
R4	도서의 출판일과 개정일은 연, 월, 일로 구성된다. 연도는 1900~2024 사이의 값을 갖고, 월과 일은 각각 01~12, 01~31 사이의 값을 갖는다.
R5	가격은 정가와 판매가(할인가)로 구분되고 양수 값을 갖는다. 가격은 원, 달러, 엔으로 표시된다. 원화인 경우 생략 가능하다(즉, 통화가 생략될 경우 원화로 간주함). 할인율(%)은 0이상 90이하의 값을 갖는다.
R6	페이지수는 1이상 1000이하의 값을 갖고, 평점(rating)은 1~10 사이의 실수 값을 갖는다.
R7	도서가 속한 분야 정보는 별도로 정의되는 도서 분류 체계에 포함된 분야(들)에 대응되어야 한다(즉, 도서 분류 체계에 존재하는 분야(들)만 이용 가능).
R8	도서의 표지그림은 외부의 이미지 파일을 지정해야 하며, 파일의 포맷(JPEG, GIF, PNG 등)과 파일을 읽고 출력할 수 있는 응용 프로그램(예: imageView.exe)을 나타내야 한다.
R9	다운로드 받을 수 있는 전자책(eBook) 도서는 위 R2에 나열된 정보들 뿐만 아니라 파일형식과 파일크기, 이용기간이 추가된다. 이용기간은 제한이 없을 경우 생략 가능하다.
R10	전자책의 파일형식은 EPUB, PDF, StreamText, StreamImage 중 하나의 값을 갖는다. 파일크기는 실수 값이고 단위가 있다(예: 985 KB, 1.45 MB 등). 이용기간은 일 수로 정의된다.
R11	새로 나온 신간 전자책(newEBook)은 이용기간이 최대 2주일이고, 개정일은 존재할 수 없다.
R12	도서 분류 체계는 문학, 인문, 사회, 과학, 유아, 컴퓨터 등 다양한 분야를 포함할 수 있다. 각 분야는 이름 뿐만 아니라 식별자를 갖고, 다시 세부 분야로 나뉠 수 있다(예: 문학은 소설, 시, 희곡 등으로 세분화될 수 있고, 소설은 다시 한국소설, 영미소설, 장르소설 등으로 더 세분화 가능). 각 분야의 이름과 분류 체계의 구성은 DTD 문서에서 미리 정의되지 않고 BookML 문서마다 사용자에게 의해 자유롭게 정의될 수 있어야 한다.

설계 방안

- R2:

- 도서는 제목, 저자, 출판사, 출판일, 가격, 페이지수, 평점, 분야에 대한 정보를 갖고, (→ 필수)
- 추가적으로 개정일, 표지그림, 책소개, 독자리뷰 등을 가질 수 있다. (→ 선택)
- 번역서인 경우 원제목과 역자 정보도 있다. (→ 선택)
- 저자와 분야, 개정일은 여러 개가 존재할 수 있고, 독자리뷰는 최대 20개까지만 가능하다. (→ + 또는 *)

→ 각 세부 정보를 도서의 속성이나 자식 엘리먼트로 정의

- R3 ~ R8을 참조하여 결정

→ 필수/선택 여부, 빈도수 정보 표시

DTD 표현 불가

x-x) 또는 ISBN13 (is

- <book id="isbn-893643412-8" ...>
<book id="isbn-978-893497246-4" ...>

<!ATTLIST book id _____>

설계 방안

- ```
<price currency="won">
 <listPrice>13000</listPrice>
 <salesPrice>11700</salesPrice>
 <discountRate>10</discountRate>
</price>
```

```
<price currency="won">
 <listPrice>13000</listPrice>
 <salesPrice
 discountRate="10">11700</salesPrice>
</price>
```

```
<price currency="won" listPrice="13000" salesPrice="11700" discountRate="10" />
```

```
<price>
 <listPrice currency="dollar">10</listPrice>
 <salesPrice currency="won" discountRate="10">11700</salesPrice>
</price>
```

```
<price>
 <listPrice currency="dollar" value="10" />
 <salesPrice currency="won" value="11700" discountRate="10" />
</price>
```

## 설계 방안

- <book id="isbn:978-893497246-4" categories="c102-1 c103-2">  
 <!-- 아래 두 분야에 속함 -->  
 ...  
</book>
- ↑ IDREF(S) 타입 속성의

```
<bookCategory>
 ...
 <category cid="c102-1" name="일반교양">... </category>
 ...
 <category cid="c103-2" name="세계사"> ... </category>
 ...
</bookCategory>
```

## 설계 방안

- R8: 도서의 표지그림은 외부의 이미지 파일을 지정해야 하며, 파일의 포맷(JPEG, GIF 등)과 이를 출력할 수 있는 응용 프로그램(예: imageView.exe)을 나타내야 한다.

→ NOTATION 타입 속성 이용 (강의자료 p.46~48 참조)

```
<coverImage src="cover.jpg" type="jpeg" />
```



```
<!NOTATION jpeg PUBLIC "image/jpeg" "imageView.exe">
<!NOTATION gif PUBLIC "image/gif" "photoViewer.exe">
...
<!ELEMENT coverImage (#PCDATA)>
<!ATTLIST coverImage src CDATA #REQUIRED
 type _____>
```

\* 주의: EMPTY element는 NOTATION 타입의 속성을 가질 수 없음.  
모든 element는 NOTATION 타입 속성을 최대 하나만 가질 수 있음

## 설계 방안

### 방법 2

```
<bookList>
 <book>
 <title>
 <author>
 ...
 <review>
 </book>
 <book>
 ...
</book>
 ...
 <eBook>
 <title>
 <author>
 ...
 <review>
 <format>
 <size>
 <period>
 </eBook>
 <eBook>
 ...
 </eBook>
 ...
</bookList>
```

다른 엘리먼트 사용

추가

## 설계 방안

- R9: 전자책(eBook) 도서는 위 R2에 나열된 도서 정보들뿐만 아니라 파일형식과 파일크기, 이용기간 정보가 추가된다.

### 방법 1

```
<bookList>
 <books>
 <book>
 <title>
 <author>
 ...
 <review>
 </book>
 </books>
 <eBooks>
 <book>
 <title>
 <author>
 ...
 <review>
 <format>
 <size>
 <period>
 </book>
 </eBooks>
</bookList>
```

동일한 엘리먼트 사용

추가

## 설계 방안

### 방법 3

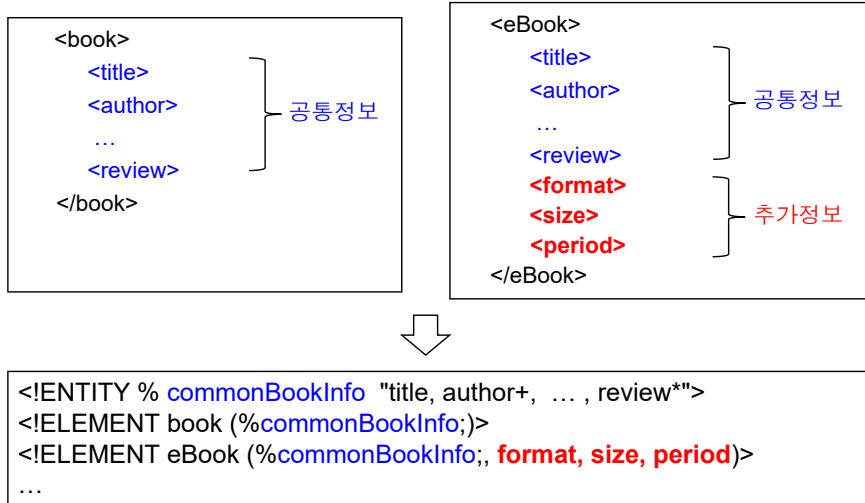
```
<bookList>
 <book type="book">
 <title>
 <author>
 ...
 <review>
 </book>
 <book type="eBook">
 <title>
 <author>
 ...
 <review>
 <format>
 <size>
 <period>
 </book>
 ...
</bookList>
```

동일한 엘리먼트와 type 속성 사용

추가

## 설계 방안

- parameter entity 활용



## 설계 방안

- R11: 새로 나온 신간 전자책(newEBook)은 이용기간이 최대 2주일 이고, 개정일은 존재할 수 없다.

```
<newEBook id="..." categories="...">
 <title>...</title>
 <author>...</author>
 ...
 <publishDate year="2024" month="02" day="15" />
 <!-- <revisionDate year="2024" month="03" day="15" /> --> <!-- 사용 불가! -->
 ...
 <format value="EPUB" />
 <size unit="MB">30.55</size>
 <period>21</period> <!-- 오류: 2주일 초과 -->
</newEBook>
```

## 설계 방안

- R10: 전자책의 파일형식은 EPUB, PDF, StreamText, StreamImage 중 하나의 값을 갖는다. 파일크기는 실수 값이고 단위가 있다 (예: 985 KB, 1.45 MB 등). 이용기간은 일 수로 정의된다.

<format value="EPUB" />	<!-- 파일형식: 미리 정해진 값(열거형) -->
<size unit="MB">26.60</size>	<!-- 파일크기: 단위는 속성으로 표현 -->
<period>14</period>	<!-- 대여기간 -->

\* 주의: 열거형 타입은 속성에만 사용 가능

<format>EPUB</format>

엘리먼트의 content는 열거형으로 정의 불가

## 설계 방안

- R12: 도서 분류 체계는 문학, 인문, 사회, 과학, 유아, 컴퓨터 등 다양한 분야를 포함할 수 있다. 각 분야는 이름뿐만 아니라 식별자(ID)를 갖고, 다시 세부 분야로 나뉠 수 있다(예: 문학은 소설, 시, 희곡 등으로 세분화될 수 있고, 소설은 다시 한국소설, 영미소설, 장르소설 등으로 더 세분화 가능). 각 분야의 이름과 분류 체계의 구성은 DTD에서 미리 정의되지 않고 BookML 문서마다 사용자에 의해 자유롭게 정의될 수 있어야 한다.

- 방법 1: 분야 별로 각각 엘리먼트 정의? → DTD에서 특정 분야들과 분류 체계를 정의

```
<bookCategory>
 <literature cid="c101">문학</literature>
 <fiction cid="c101-1">소설</fiction>
 ...
 <humanities cid="c102">인문</humanities>
 ...
</bookCategory>
```

## 설계 방안

- 방법 2: 모든 분야에 대해 하나의 공통 엘리먼트 사용
  - 모든 분야는 동일한 수준(level)

```
<bookCategory>
 <category cid="c101" name="문학" />
 <category cid="c101-1" name="소설" />
 ...
 <category cid="c102" name="인문" />
 ...
</bookCategory>
```

## 설계 방안

- 방법 3: 계층적/재귀적 구조 (recursive element)
  - <category>가 다른 <category>를 자식 엘리먼트로 가질 수 있도록 정의 → 가장 좋은 방법

```
<bookCategory>
 <category cid="c101" name="문학">
 <category cid="c101-1" name="소설">
 <category cid="c101-1-1" name="한국소설"> ... </category>
 <category cid="c101-1-2" name="영미소설"> ... </category>
 ...
 </category>
 <category cid="c101-2" name="시"> ... </category>
 <category cid="c101-3" name="희곡"> ... </category>
 ...
</category>
<category cid="c102" name="인문">
 <category cid="c102-1" name="일반교양"> ... </category>
 ...
</category>
```