

XML의 이해

김 현

한국학중앙연구원 인문정보학교실

hyeon@aks.ac.kr



이 저작물(PPT)의 인용 표시 방법:
김현, 「XML의 이해」, '전자문서와 하이퍼텍스트' 수업 자료 (2020)

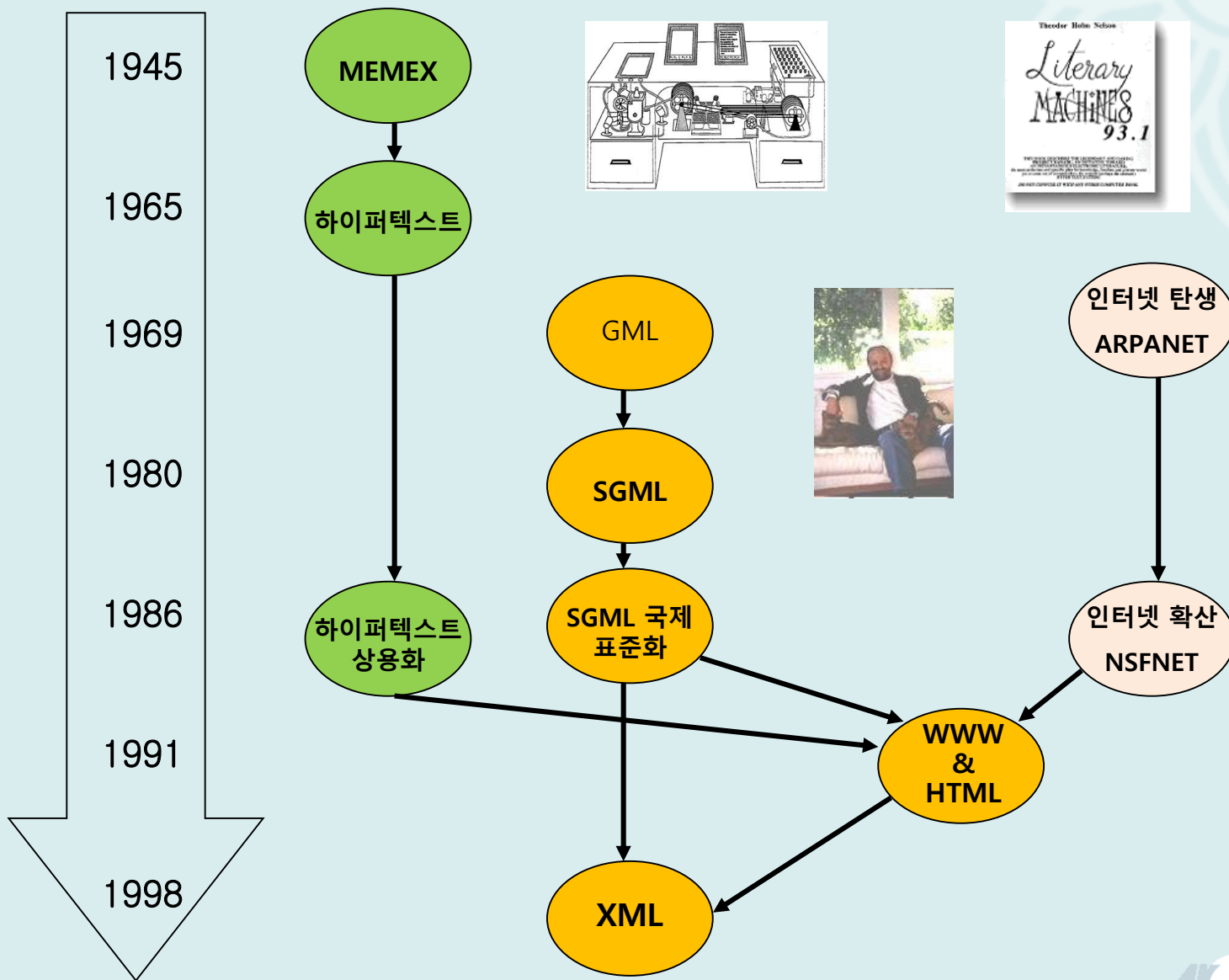


1. XML 발전의 역사

2. XML의 특징

3. Namespace의 활용

- ❖ XML: eXtensible Markup Language
 - 전자문서 기술 언어의 세계적인 표준(W3C-recommended general-purpose markup language.)
 - W3C Recommendation: XML 문서의 구조와 XML 프로세서의 요건을 정의
- ❖ General-Purpose
 - 특정 분야, 특정 목적의 전자문서 생산에 적용할 수 있는 보편적 표준을 제공
- ❖ Interoperable
 - 시스템의 상이함에 구애되지 않는 데이터 공유
- ❖ Open, Free Standard
 - 누구나 사용할 수 있는 개방된 표준



SGML (Standard Generalized Markup Language)

- ❖ 문서에 담긴 정보 요소와 그 요소들간의 관계를 컴퓨터가 이해할 수 있도록 만든 마크업 언어
- ❖ 1969년 미국의 IBM 사의 찰스 골드파브(Charles Goldfarb)와 그 동료들에 의해 GML(Generalized Markup Language)이라는 이름으로 처음 개발
- ❖ 그 후 지속적으로 발전하여 1986년 국제표준기구에 의해 구조적 데이터 표현의 표준안(ISO 8876)으로 인증



Fathers of **GML**

- Charles **G**oldfarb
- Ed **M**osher
- Ray **L**orie

HTML (Hypertext Markup Language)

- ❖ HTML도 기본적으로는 SGML을 기반으로 만들어진 것.
- ❖ 그러나 HTML은 문서를 '내용', '구조', '모양'으로 구분하여 그 각각의 독립성과 유연성을 추구한 SGML의 기본 사상에 충실하지 않았다.
- ❖ 대신 누구나 쉽게 '하이퍼텍스트' 문서를 만들 수 있도록 하였고, 그것에 의해 월드 와이드 웹이 탄생.
- ❖ SGML을 차용하되, SGML에 얽매이지 않는 것이 HTML의 성공 요인

XML (eXtensible Markup Language)

- ❖ HTML이 약화시켰던 SGML의 중요한 목적을 다시 회복
 - ☞ HTML이 표현하지 못했던 문서의 구조와 내용 요소의 성격을 명확하게 표현
- ❖ XML 문서의 대표적 특성: 문서를 “문서의 구조”, “문서의 내용”, “문서의 모양” 3 가지로 분리한다는 것.
 - ☞ 문서의 내용과 모양을 구분함으로써, 내용을 그대로 두고 모양만을 바꾼다거나, 모양에 손대지 않은 채 내용을 확장하고 수정할 수 있는 유연성 보장
- ❖ XML은 인터넷과 하이퍼텍스트가 결합한 곳에서 텍스트의 구조와 정보 요소까지 드러내고자 하는 취지에서 개발된 것.



1. XML 발전의 역사

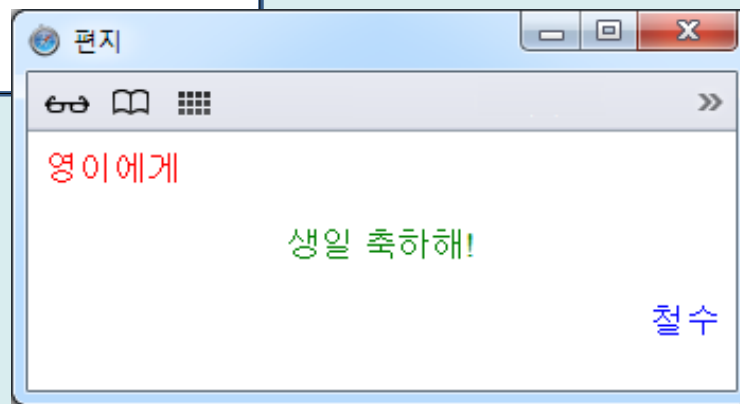
2. XML의 특징

3. Namespace의 활용

HTML:구조,내용,모양의 통합

HTML: 문서의 내용과 모양을 한꺼번에 취급

```
<html>
  <head>
    <title>편지</title>
  </head>
  <body>
    <p style="color:red; text-align:left">영이에게</p>
    <p style="color:green; text-align:center">생일축하해!</p>
    <p style="color:blue; text-align:right">철수</p>
  </body>
</html>
```



XML: 구조, 내용, 모양의 분리

XSD: 문서의 구조

```
<?xml version="1.0" encoding="utf-8"?>

<xs:schema targetNamespace="letter.xsd"
  xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <xs:element name="편지">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="수신"/>
        <xs:element ref="본문"/>
        <xs:element ref="발신"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>

  <xs:element name="수신" type="xs:string"/>
  <xs:element name="본문" type="xs:string"/>
  <xs:element name="발신" type="xs:string"/>

</xs:schema>
```

XML: 문서의 내용

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/xsl" href="letter.xsl"?>

<편지 xmlns="letter.xsd">
  <수신>영이에게</수신>
  <본문>생일축하해</본문>
  <발신>철수</발신>
</편지>
```

XSL: 문서의 모양

```
<?xml version="1.0" encoding="utf-8"?>

<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:ltt="letter.xsd">

  <xsl:template match="/">
    <html>
      <head>
        <title>편지</title>
      </head>
      <body>
        <xsl:apply-templates/>
      </body>
    </html>
  </xsl:template>

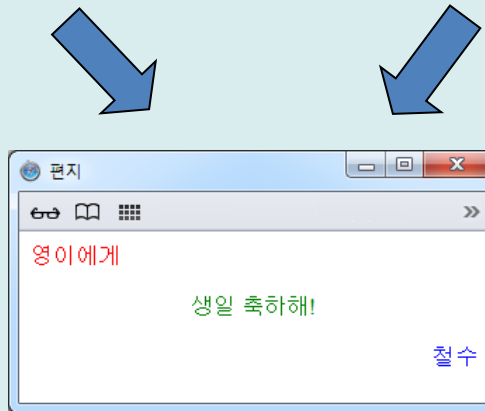
  <xsl:template match="ltt:편지">
    <xsl:apply-templates/>
  </xsl:template>

  <xsl:template match="ltt:수신">
    <p style="color:red; text-align:left">
      <xsl:value-of select="."/> </p>
  </xsl:template>

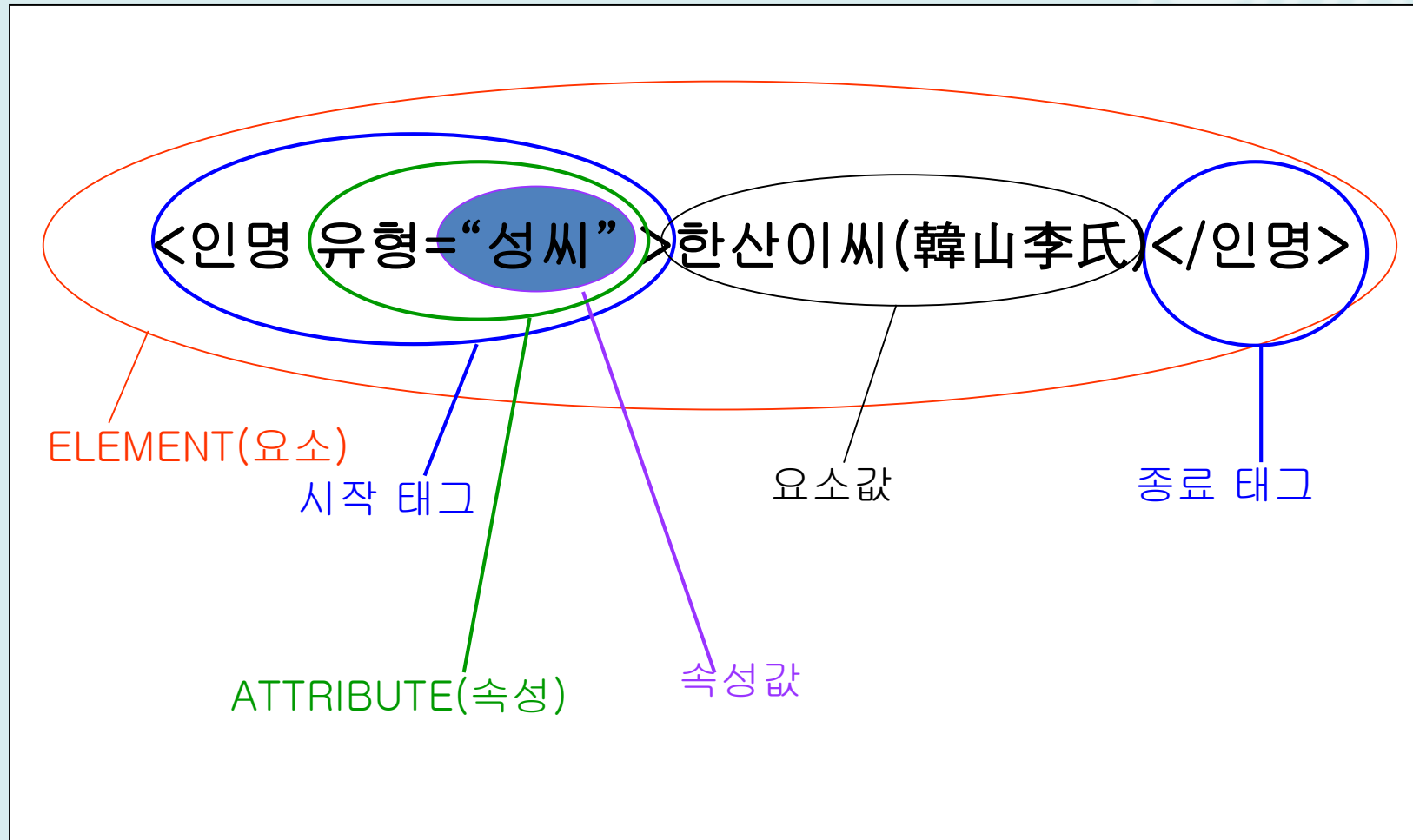
  <xsl:template match="ltt:본문">
    <p style="color:green; text-align:center">
      <xsl:value-of select="."/> </p>
  </xsl:template>

  <xsl:template match="ltt:발신">
    <p style="color:blue; text-align:right">
      <xsl:value-of select="."/> </p>
  </xsl:template>

</xsl:stylesheet>
```



XML의 구성 요소



XML 문서 구조

```
<?xml version="1.0" encoding="utf-8"?>
```

```
<?xml-stylesheet type="text/xsl" href="수내동가옥.xsl"?>
```

```
<항목 ID="10001932">
```

```
<항목명>
```

```
수내동 가옥
```

```
</항목명>
```

```
<메타데이터>
```

```
<분야>생활·민속</분야>
```

```
<유형>건물</유형>
```

```
</메타데이터>
```

```
<본문>
```

```
<지명>분당</지명>신도시 개발 전 이곳에 세거하였던 <인명 유형="성씨">한산  
이씨(韓山李氏)</인명>의 살림집 중 한 채로 <인명>이택구</인명>씨가 거주하였던 집이다.
```

```
</본문>
```

```
</항목>
```

구조 요소

최상위 요소

문중 요소

XSD (XML Schema Definition) 예시

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<xs:schema
xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
```

```
<xs:element name="항목">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="항목명"/>
      <xs:element ref="메타데이터"/>
      <xs:element ref="본문"/>
    </xs:sequence>
    <xs:attributeGroup ref="attlist.항목"/>
  </xs:complexType>
</xs:element>
```

요소 선언

```
<xs:attributeGroup name="attlist.항목">
  <xs:attribute name="ID" use="required"/>
</xs:attributeGroup>
```

속성 선언

```
<xs:element name="항목명" type="xs:string"/>
```

```
<xs:element name="메타데이터">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="분야"/>
      <xs:element ref="유형"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

```
<xs:element name="분야" type="xs:string"/>
```

```
<xs:element name="유형" type="xs:string"/>
```

```
<xs:element name="본문">
  <xs:complexType mixed="true">
    <xs:choice minOccurs="0" maxOccurs="unbounded">
      <xs:element ref="인명"/>
      <xs:element ref="지명"/>
    </xs:choice>
  </xs:complexType>
</xs:element>
```

```
<xs:element name="인명" type="xs:string"/>
```

```
<xs:attributeGroup name="attlist.인명">
  <xs:attribute name="유형"/>
</xs:attributeGroup>
```

```
<xs:element name="지명" type="xs:string"/>
```

요소 선언

```
<xs:attributeGroup name="attlist.지명">
  <xs:attribute name="유형"/>
</xs:attributeGroup>
```

속성 선언

```
</xs:schema>
```

DTD (Document Type Definition) 예시

```
<?xml version="1.0" encoding="utf-8"?>
```

```
<!ENTITY % 고유명사 "인명|지명">
```

개체(ENTITY) 정의

```
<!ELEMENT 항목 (항목명, 메타데이터, 본문)>
```

요소 선언

```
<!ATTLIST 항목 ID CDATA #REQUIRED>
```

```
<!ELEMENT 항목명 (#PCDATA)>
```

```
<!ELEMENT 메타데이터 (분야, 유형)>
```

```
<!ELEMENT 분야 (#PCDATA)>
```

```
<!ELEMENT 유형 (#PCDATA)>
```

```
<!ELEMENT 본문 (#PCDATA|%고유명사;)*>
```

```
<!ELEMENT 인명 (#PCDATA)>
```

```
<!ATTLIST 인명 유형 CDATA #IMPLIED>
```

```
<!ELEMENT 지명 (#PCDATA)>
```

```
<!ATTLIST 지명 유형 CDATA #IMPLIED>
```

속성 선언

XSL (Extensible Stylesheet Language) 예시(1)

```
<?xml version="1.0" encoding="utf-8"?>  
<!-- XSL for 항목_Display 1.0 written by Hyeon Kim 2014. 5. 8. -->
```

XSL 문서 최상위 요소

```
<xsl:stylesheet version="1.0"  
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"  
  xmlns:html="http://www.w3.org/1999/xhtml">
```

```
<xsl:output method="xml" indent="yes"/>
```

```
<xsl:template match="/">  
  <html xmlns="http://www.w3.org/1999/xhtml">
```

```
    <head>
```

```
      <title>Site Scene</title>
```

```
      <style>
```

```
        body {font-family:함초롬바탕}
```

```
        img {border:0}
```

```
        a {text-decoration:none}
```

```
      </style>
```

```
    </head>
```

```
    <body>
```

```
      <xsl:apply-templates/>
```

```
    </body>
```

```
  </html>
```

```
</xsl:template>
```

Namespace for XSL
Elements

Namespace for HTML
Elements

XPATH

XSL (Extensible Stylesheet Language) 예시(2)

```

<xsl:template match="항목">
  <xsl:apply-templates />
</xsl:template>

<xsl:template match="항목명">
  <html:font
    style="color:red;font-weight:bold;font-size:18pt">
    <xsl:value-of select="."/>
  </html:font>

<html:br><br></html:br><html:br><br></html:br>
</xsl:template>

<xsl:template match="메타데이터">
  <xsl:apply-templates/>
  <html:br><br></html:br>
</xsl:template>

<xsl:template match="분야">
  <html:font style="color:grey;font-weight:bold">
o 분야: </html:font>
    <xsl:value-of select="."/>
    <html:br><br></html:br>
  </xsl:template>

<xsl:template match="유형">
  <html:font style="color:grey;font-weight:bold">
o 유형: </html:font>
    <xsl:value-of select="."/>
    <html:br><br></html:br>
  </xsl:template>

```

```

<xsl:template match="본문">
  <xsl:apply-templates/>
  <html:br><br></html:br>
</xsl:template>

<xsl:template match="지명">
  <html:font style="color:blue;font-weight:bold">
    <xsl:value-of select="."/>
  </html:font>
</xsl:template>

<xsl:template match="인명">
  <html:font style="color:green;font-weight:bold">
    <xsl:value-of select="."/>
  </html:font>
</xsl:template>

<xsl:template match="text()">
  <xsl:value-of select="."/>
</xsl:template>

</xsl:stylesheet>

```

XML 관련 기술

- ❖ XSLT(Extensible Stylesheet Language Transformation): XML 문서의 내용을 컴퓨터상에서 시각적으로 표현하기 위한 기술.
 - XML 구조를 HTML이나 다른 언어의 구조로 변환하는 기능을 지원.
- ❖ XPath(XML Path Language): XML 문서 내의 특정 부분을 찾아가기 위해 사용되는 경로 표기 언어.
- ❖ DOM(Document Object Model): XML 문서를 컴퓨터가 구조적인 정보로 취급할 수 있게 하는 기술.
- ❖ XML RDF (Resource Description Framework): RDF는 월드 와이드 웹의 자원을 컴퓨터가 인식할 수 있도록 하는 기술적 체계. XML RDF는 이것을 XML 문서 형태로 기술하는 방법을 제공



1. XML 발전의 역사

2. XML의 특징

3. Namespace의 활용

Namespace의 정의

❖ Namespace란?

- XML 문서 내에서 일련의 요소와 속성 이름을 명확하게 인식할 수 있게 하는 식별자
- 하나의 그룹을 이루는 요소 또는 속성 이름의 집합
- A unique identifier used to group a set of XML names (elements or attributes). Namespaces provide a way to qualify element and attribute names so that names in schemas do not conflict.

❖ Namespace의 필요성

- XML 문서 상에서는 성격이 다른 2개 이상의 요소가 동일한 이름을 갖는 경우도 발생할 수 있음.
- 각각의 요소가 각기 다른 영역에 속하는 것임을 표시해 주는 장치가 필요하다.

Namespace가 필요한 경우

❖ 요소 이름의 중복 또는 내용 및 구조의 불일치

<고문헌해제>

<해제 번호="101">

<제목>유서(諭書)</제목>

<서지사항>

<작성연도>.....</작성연도>

<크기>.....</크기>

</서지사항>

<내용>

<정의>肅宗이 林泳에게 臘劑와 柴炭을 하사하면서 내린 諭書. </정의>

<주제>본 문서는 御札諭書帖에 있는 諭書 중 한 가지인데,</주제>

<인물>林泳(1649-1696)은 본관이 羅州,</인물>

</내용>

<참고문헌>『大典會通』,『古文書集成』 32 -慶州慶州孫氏篇-, 韓國精神文化研究院, 2002.</참고문헌>

<이미지>유서.bmp</이미지>

</해제>

<해제 번호="201">

<제목>신식유서필지(新式儒胥必知)</제목>

<저자>황필수(黃泌秀)</저자>

<간행연대>광무 5년(1901)</간행연대>

<권수>1권 1책 43장</권수>

<크기>.....</크기>

<내용>총래의 「유서필지」를 저본으로 하여 갑오개혁 이후의 변화한 사항 등을 개정하여 새롭게 간행</고도서:내용>

<이미지>신식유서필지.jpg</이미지>

</해제>

</고문헌해제>

Namespace의 선언

❖ 기본 Namespace의 선언

```
<element_name xmlns="URI_Reference">
```

❖ 접두사를 사용한 Namespace의 선언

```
<prefix:element_name xmlns:prefix="URI_Reference">
```

Namespace의 활용 예시

```

<고문헌해제 xmlns:고문서="http://digerati.aks.ac.kr/terms/문헌/고문서해제"
             xmlns:고도서="http://digerati.aks.ac.kr/terms/문헌/고서해제" >

  <고문서:해제 번호="101">
    <고문서:제목>유서(諭書)</고문서:제목>
    <고문서:서지사항>
      <고문서:작성연도>.....</고문서:작성연도>
      <고문서:크기>.....</고문서:크기>
    </고문서:서지사항>
    <고문서:내용>
      <고문서:정의>肅宗이 林泳에게 臘劑와 柴炭을 하사하면서 내린 諭書. </고문서:정의>
      <고문서:주제>본 문서는 御札諭書帖에 있는 諭書 중 한 가지인데, .....</고문서:주제>
      <고문서:인물>林泳(1649-1696)은 본관이 羅州, .....</고문서:인물>
    </고문서:내용>
    <고문서:참고문헌>『大典會通』, 『古文書集成』 32 -慶州慶州孫氏篇-, 韓國精神文化研究院, 2002.</고문서:참고문헌>
    <고문서:이미지>유서.bmp</고문서:이미지>
  </고문서:해제>

  <고도서:해제 번호="201">
    <고도서:제목>신식유서필지(新式儒胥必知)</고도서:제목>
    <고도서:저자> 황필수(黃泌秀)</고도서:저자>
    <고도서:간행연대>광무 5년(1901)</고도서:간행연대>
    <고도서:권수>1권 1책 43장</고도서:권수>
    <고도서:크기>.....</고도서:크기>
    <고도서:내용>
      종래의 「유서필지」를 저본으로 하여 갑오개혁 이후의 변화한 사항 등을 개정하여 새롭게 간행
    </고도서:내용>
    <고도서:이미지>신식유서필지.jpg</고도서:이미지>
  </고도서:해제>

</고문헌해제>

```