

Ch. 3 建立XML文件

- XML文件的基本結構
- 編輯XML文件
- XML文件的宣告
- XML文件的語法
- XML文件的屬性
- XML文件的名稱空間
- 建立XHTML文件

1

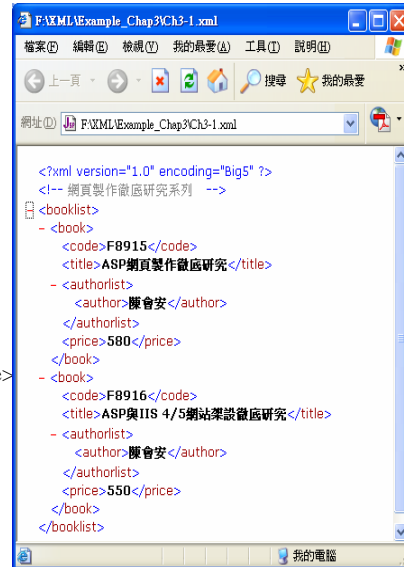
XML文件的基本結構

- 一份XML文件
- XML文件的組成
- 使用IE瀏覽XML文件
- XML文件的結構

2

一份XML文件(Ch3-1.xml)

```
01: <?xml version="1.0" encoding="Big5"?>
02: <!--網頁製作徹底研究系列-->
03: <booklist>
04:   <book>
05:     <code>F8915</code>
06:     <title>ASP網頁製作徹底研究</title>
07:     <authorlist>
08:       <author>陳會安</author>
09:     </authorlist>
10:     <price>580</price>
11:   </book>
12:   <book>
13:     <code>F8916</code>
14:     <title>ASP與IIS 4/5網站架設徹底研究</title>
15:     <authorlist>
16:       <author>陳會安</author>
17:     </authorlist>
18:     <price>550</price>
19:   </book>
20: </booklist>
```



一份XML文件

- XML文件的基本架構，可以分為下列幾個部分，如下所示：
 - 文件宣告：第1列就是XML文件的宣告，它定義XML文件的版本和使用的字碼集，以此例為1.0版，使用中文Big5字碼。
 - 根標籤：第3列為XML文件的根標籤 <booklist>，這也就是樹狀結構的根節點。
 - 子元素：在第4~19列就是根元素的子元素 book，而第5~10列和第13~18列為book元素的子元素code、title、authorlist和price。

XML文件的組成

元素(Element)

- XML文件是由元素所組成，XML的元素和標籤是不同的，其代表的意義，如下所示：
 - 標籤(Tag)：XML能夠自己定義標籤，一個標籤是用來標示文件的部分內容，例如：標籤<code>、<title>和<price>等，標籤分為開頭標籤<code>和結尾標籤</code>。
 - 元素(Element)：XML元素為整個文件的主要架構，元素的本身可以是標籤加上文字內容，或是元素內包含有其它的元素，元素是一個完整的項目，它包含標籤、屬性、開始標籤和結尾標籤內的文字內容和結尾標籤。

5

XML文件的組成

元素(Element)

- XML元素是由開始和結尾標籤構成，其中包含文字內容，如下所示：

```
<code>F8915</code>
```

- 此外XML元素允許包含其它的子元素，如下所示：

```
<book>
```

```
  <code>F8915</code>
```

```
  <title>ASP網頁製作徹底研究</title>
```

```
  <authorlist>
```

```
    <author>陳會安</author>
```

```
  </authorlist>
```

```
  <price>580</price>
```

```
</book>
```

6

XML文件的組成

屬性(Attribute)

- 在開始標籤的標籤名稱後，可以替標籤加上屬性名稱，每一個屬性都擁有屬性名稱和值，而且值需要使用引號括起，如下所示：

```
<book code="F8915">
```

- 上述<book>標籤擁有code屬性。

7

XML文件的組成

實體參考(Entity Reference)

- 標籤語言本身擁有一些保留符號，例如：標籤中的「<」符號，如果文件內容需要使用到這些符號，此時需要使用實體參考。
- 在XML有五個預先保留的實體參考，如下表所示：

實體參考	符號
<	<
>	>
&	&
'	'
"	"

8

XML文件的組成

實體參考(Entity Reference)

- 每一個實體參考都是由「&」符號開始，然後以「;」結束，例如：XML文件不可以寫成下列的文件內容，如下所示：

```
<order>書價 < 500</order>
```

- 上述的標籤內容有「<」符號，我們需要使用實體參考，如下所示：

```
<order>書價 &lt; 500</order>
```

9

XML文件的組成

註解(Comment)

- XML的註解和HTML是一樣的，都是由「<!--」開始和「-->」結尾，如下所示：

```
<!--網頁製作徹底研究系列-->
```
- 基本上註解並不屬於XML文件的內容，XML剖析器並不會處理註解文字。

10

XML文件的組成

CDATA區塊(CDATA Section)與PCDATA

- 在XML文件的文字內容就是PCDATA，當在處理XML元素title時，標籤中的文字內容【ASP網頁製作徹底研究】就是PCDATA，如下所示：

```
<title>ASP網頁製作徹底研究</title>
```

11

XML文件的組成

CDATA區塊(CDATA Section)與PCDATA

- CDATA區塊告訴XML剖析器不用處理此區塊的內容，它是位在「<![CDATA[」字串和「]]>」字串間的文字內容，如下所示：

```
<script>  
<![CDATA[  
  function add(a, b)  
  {  
    return a+b  
  }  
]]>  
</script>
```

12

XML文件的組成

PI(Processing Instructions)

- PI允許XML文件包含應用程式的指令，例如：在XML文件的開頭宣告就是一個PI，這些資訊告訴XML剖析器如何處理XML文件，如下所示：

```
<?xml version="1.0" encoding="Big5"?>
```

- 上述PI如同註解一般並非XML文件的內容，不過不同註解，XML剖析器會將這些資料送給應用程式。

13

XML文件的組成

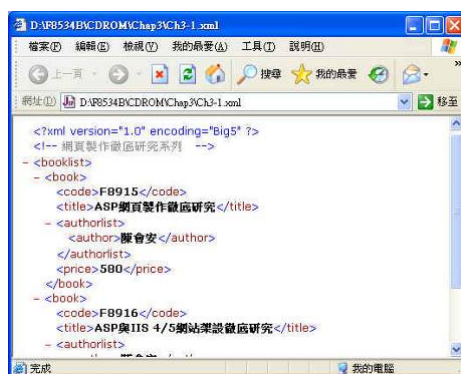
DTD(Document Type Declarations)

- XML規格有很大的部分屬於DTD文件驗證指令，這是SGML預設的文件驗證方式，也被XML所採用，在XML文件除了元素外，就是檢查這些元素的DTD。

14

使用IE瀏覽XML文件

- 對於XML文件我們可以如同HTML文件一般，使用Internet Explorer瀏覽，請直接按二下XML文件或執行「檔案/開啟舊檔」指令開啟XML文件，如下圖所示：



15

XML文件的結構

- XML文件的結構是由XML元素組成，在內容部分並沒有一定的文件結構，因為標籤為使用者自行定義，在XML文件的元素定義方式將產生不同的文件結構，例如：原來的book元素，如下所示：

```
<book>
  <code>F8915</code>
  <title>ASP網頁製作徹底研究</title>
  <authorlist>
    <author>陳會安</author>
  </authorlist>
  <price>580</price>
</book>
```

16

XML文件的結構

- authorlist子元素author可以再分為姓和名的子元素，如下所示：

```
<book>
  <code>F8915</code>
  <title>ASP網頁製作徹底研究</title>
  <authorlist>
    <author>
      <lastname>陳</lastname>
      <fristname>會安</fristname>
    </author>
  </authorlist>
  <price>580</price>
</book>
```

17

XML文件的結構

- 轉換成樹狀結構，馬上可以看出其間的差異，這個文件的樹高比原來的高，因為擁有**更多的階層**。
- 至於文件結構是否需要分的如此細，全憑應用程式的需求，例如：如果應用程式需要處理姓的話，就需要將author元素進一步分成姓和名，反之如果只需姓名，那直接使用author元素即可。

18

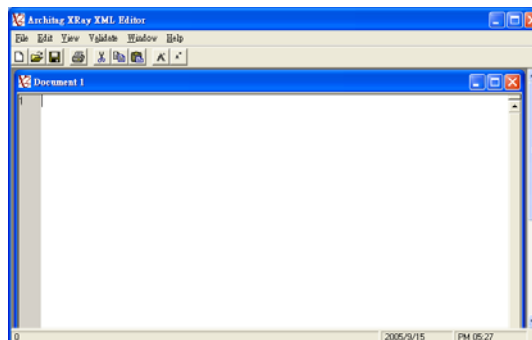
編輯XML文件

- 啟動 XRay XML Editor
- 建立與編輯 XML 文件
- 儲存 XML 文件

19

啟動 XRay XML Editor

- 在 Windows XP 作業系統安裝 XRay XML Editor 後，就可以執行「開始/所有程式/Architag XRay XML Editor/XRay XML Editor」指令來啟動 XRay XML Editor，如下圖所示：



20

建立與編輯 XML 文件

- XRay 是一套 XML 學習工具，可以在輸入 XML 文件時，馬上即時回應 XML 文件是否是一份良好格式的文件，換句話說，每一次的鍵盤輸入，XRay 就會即時剖析和檢查是否有錯誤，如果有任何錯誤，馬上顯示錯誤訊息來幫助使用者建立良好格式的 XML 文件。

21

建立與編輯 XML 文件

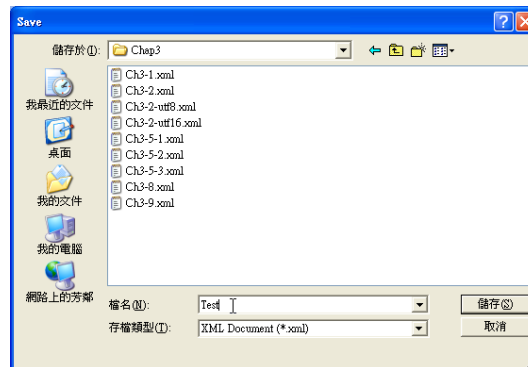
```
Document 3
1 <?xml version="1.0" encoding="Big5"?>
2 <!--網頁製作徹底研究系列-->
3 <booklist>
4 <book code="F8915">
5 <title>ASP網頁製作徹底研究</title>
6 <authorlist>
7 <author>陳會安</author>
8 </authorlist>
9 <price>580</price>
10 </book>
11 <book code="F8916">
12 <title>ASP與IIS 4/5網站架設徹底研究</title>
13 <authorlist>
14 <author>陳會安</author>
15 </authorlist>
16 <price>580</price>
17 </book>
18 </booklist>
```

```
Document 3
1 <?xml version="1.0" encoding="Big5"?>
2 <!--網頁製作徹底研究系列-->
3 <booklist>
4 <book code="F8915">
5 <title>ASP網頁製作徹底研究</title|
6 <authorlist>
7 <author>陳會安</author>
8 </authorlist>
9 <price>580</price>
10 </book>
11 <book code="F8916">
12 <title>ASP與IIS 4/5網站架設徹底研究</title>
13 <authorlist>
```

Line: 5, Character: 27
<title>ASP網頁製作徹底研究</title
* End tag 'title' does not match the start tag 'title'. (-1072896659) error Ln 5, Col 32

儲存 XML 文件

- 在 XRay 完成 XML 文件編輯後，請執行「File/Save」指令儲存 XML 文件，可以看到「Save」對話方塊。



23

XML文件的宣告

- 版本屬性version
- 字碼集屬性encoding
- 獨立文件屬性standalone

24

版本屬性version

- 在XML文件的宣告部分，最主要的屬性就是XML的版本，目前XML的版本為1.0版，如下所示：

```
<?xml version="1.0"?>
```

- 上述屬性說明此文件的版本為1.0版的規格，這是告訴XML剖析器使用那個版本的XML規格處理XML文件的內容。

25

字碼集屬性encoding

- XML文件可能包含各國文字的編碼，例如：東方語系的雙位元的字元，此時我們需要指定字碼集，才能正確的在瀏覽程式顯示文件的內容，繁體中文就是Big5，如下所示：

```
<?xml version="1.0" encoding="Big5"?>
```

- 雖然Internet Explorer 5.x瀏覽程式支援Big5，但是並不是所有的XML工具都支援Big5，此時可以使用UTF-8和UTF-16的「統一字碼」(Unicode)。

26

字碼集屬性encoding

統一字碼(Unicode)

- 統一字碼是由Unicode Consortium組織所制定的一個能包括全世界文字的字碼集，它包含GB2312和Big5字碼集的所有字集，即ISO 10646字集，它有兩種編碼方式：UTF-8為8位元編碼，UTF-16為16位元的編碼。

27

獨立文件屬性standalone

- 在使用XML剖析器時，如果XML文件本身已經擁有DTD，但是也同時使用外部的DTD檔案，使用standalone屬性設定是否同時使用外部的DTD檔案，還是只使用內部的DTD，如下所示：

```
<?xml version="1.0" encoding="Big5"  
standalone="no"?>
```

- Standalone的屬性值只有兩種，如下所示：
 - Yes：不使用外部的DTD檔案。
 - No：使用外部的DTD檔案。

28

XML文件的語法

- XML標籤
- XML元素
- 建立XML文件

29

XML標籤

- 在XML文件的元素都是使用標籤標示，XML標籤可以分為：
 - 開始標籤(Start-tag)
 - 結尾標籤(End-tag)
 - 空元素標籤(Empty-element Tags)

30

XML標籤

開始標籤的語法

`<tag_name attribute_name="attribute_value">`

- 上述語法的標籤擁有屬性，其各參數的說明如下所示：
 - tag_name：XML元素的名稱，如果為英文，大小寫是不同的標籤。
 - attribute_name：屬性名稱。
 - attribute_value：屬性值，需要使用單或雙引號括起，如果不只一個，請使用空白分隔。

31

XML標籤

結尾標籤的語法

`</tag_name>`

- 上述tag_name為結尾標籤的名稱，這個名稱需要和開始標籤相同，而且成雙成對的出現，如下所示：

`</code>`

32

XML標籤

空元素(Empty Element)標籤的語法

- 如果開始標籤和結尾標籤中並沒有任何的內容，我們稱為空元素標籤，它有兩種寫法，如下所示：
`<tag_name />`
`<tag_name></tag_name>`
- 上述兩種寫法都是空元素標籤，如果空元素標籤沒有結尾標籤，標籤需要使用「/>」符號結束。

33

XML元素

XML元素的語法

- `<tag_name>content</tag_name>`
- 上述XML元素使用標籤包含內容content，content就是元素的內容，它可能包含其它的XML元素，如下所示：
`<code>F8915</code>`
- 元素內容也可以擁有其它的元素，例如：book元素，如下所示：
`<book>`
`<code>F8915</code>`
`<title>ASP網頁製作徹底研究</title>`
`</book>`
- XML元素中也可以有文字字串的內容
`<para>ASP網頁製作徹底研究</parp>`

34

建立XML文件

- XML文件需要有根標籤

```
<root>  
  <child>  
    <sub_child> .... </sub_child>  
  </child>  
</root>
```

- XML的標籤大小寫不同，可以使用中文
- 所有元素需要有結尾標籤
- 巢狀XML元素不可以重疊
- 屬性值需要有引號

35

XML文件的屬性

- XML元素的屬性
- 正確的使用元素和屬性

36

XML元素的屬性(Ch3-5-1.xml)

01: <?xml version="1.0" encoding="Big5"?>
02: <!--網頁製作徹底研究系列-->
03: <booklist>
04: <book id="1">
05: <code>F8915</code>
06: <title>ASP網頁製作徹底研究</title>
07: <authorlist>
08: <author>陳會安</author>
09: </authorlist>
10: <price>580</price>
11: </book>

37

XML元素的屬性

12: <book id="2">
13: <code>F8916</code>
14: <title>ASP與IIS 4/5網站架設徹底研究</title>
15: <authorlist>
16: <author>陳會安</author>
17: </authorlist>
18: <price>550</price>
19: </book>
20: </booklist>

38

正確的使用元素和屬性

- 屬性值不能包含多個值，但是元素可以。
- 屬性無法描述文件的結構，元素可以建立文件的樹狀結構。
- 屬性在維護上和擴充上不如元素，過多的屬性將造成維護上的困擾。
- 元素能夠建立成樹狀結構，在程式設計上比使用屬性在處理上更為簡單。

39

XML文件的名稱空間

- 什麼是名稱空間
- 名稱空間的宣告
- 名稱空間的範圍

40

什麼是名稱空間

- 「名稱空間」(Namespace)是為了解決標籤名稱重複的問題，它提供標籤或屬性更進一步的辨識資訊。

- XML文件標籤和屬性的完整名稱，如下所示：

<名稱空間字頭:標籤名稱>

- 標籤完整名稱分為名稱空間的「字頭名稱」(Prefix)，這是用來代替名稱空間字串「URI」(Uniform Resource Identifier)，和標籤名稱，雖然後面標籤的名稱相同，只要前面的名稱空間不同，它們就屬於不同的標籤，如下所示：

<bk:title>

<jb:title>

41

名稱空間的宣告

- 名稱空間的語法稱為URI，這是一個字串用來識別抽象或實際的資源，例如：網頁或下載檔案的網址。基本上URI擁有兩種格式，如下所示：

- URL(Uniform Resource Locator)：這是使用網址名稱和路徑，因為公司網址的網域名稱是唯一的，換句話說，名稱空間定義的標籤，就成為獨一無二的標籤，如下所示：

<http://www.company.com/>

<http://www.company.com/xml/>

- URN(Uniform Resource Names)：URN宣告的名稱空間的格式，如下所示：

urn:<名稱空間識別>:<名稱空間的特殊字串>

urn:company.com:book

urn:www.takming.edu.tw:school

42

名稱空間的宣告

預設宣告(Default Declaration)

- 在同一個XML元素的範圍內，宣告預設的名稱空間，只需在根元素宣告名稱空間，其格式如下所示：

```
xmlns:prefix="URI"
```

- 上述指令格式在屬性xmlns後額外加上「:」號後的字頭名稱，在XML文件根元素使用字頭名稱代替名稱空間，其子元素都屬於同一個名稱空間。

43

名稱空間的宣告

預設宣告(Default Declaration)

```
<bk:book xmlns:bk="http://www.company.com/">  
  <code>F8915</code>  
  <title>ASP網頁製作徹底研究</title>  
  <authorlist>  
    <author>陳會安</author>  
  </authorlist>  
  <price currency="NT">580</price>  
</bk:book>
```

44

名稱空間的宣告

明確宣告(Explicit Declaration)

- 如果在XML文件中元素的子元素屬於不同的名稱空間，我們可以使用字頭名稱明確宣告不同的名稱空間，在XML文件的標籤使用此字頭名稱代替名稱空間，如下所示：

```
<bk:book xmlns:bk="http://www.company.com/"
  xmlns:jb="http://www.flag.com.tw/">
  <bk:code>F8915</bk:code>
  <bk:title>ASP網頁製作徹底研究</bk:title>
```

45

名稱空間的宣告

明確宣告(Explicit Declaration)

```
<bk:authorlist>
  <bk:author>陳會安</bk:author>
  <jb:job>
    <jb:companyname>旗標出版股份有限公司
  </jb:companyname>
  <jb:title>作者</jb:title>
</jb:job>
</bk:authorlist>
<bk:price bk:currency="NT">580</bk:price>
</bk:book>
```

46

名稱空間的範圍

```

<booklist>
  <bk:book xmlns:bk="http://www.company.com/">
    <code>F8915</code>
    <title>ASP網頁製作徹底研究</title>
    <authorlist>
      <author>陳會安</author>
      <jb:job xmlns:jb="http://www.flag.com.tw/">
        <companyname>旗標出版股份有限公司</companyname>
        <title>作者</title>
      </jb:job>
    </authorlist>
    <price currency="NT">580</price>
  </bk:book>
</booklist>

```

47

名稱空間的範圍-2

元素或屬性	元素或屬性名稱	名稱空間
元素	booklist	
元素	book	http://www.company.com/
屬性	xmlns	
元素	code	http://www.company.com/
元素	title	http://www.company.com/
元素	authorlist	http://www.company.com/
元素	author	http://www.company.com/
元素	job	http://www.flag.com.tw/
屬性	xmlns	
元素	companyname	http://www.flag.com.tw/
元素	title	http://www.flag.com.tw/
元素	price	http://www.company.com/
屬性	currency	http://www.company.com/

48

建立XHTML文件

- 什麼是XHTML文件
- XHTML文件的語法
- XHTML文件的DTD
- XHTML就是良好格式的HTML

49

什麼是XHTML文件

- 「XHTML」(eXtensible HyperText Markup Language)規格是在2000年1月才正式通過，雖然規格比XML 1.0版還晚制定，事實上它就是HTML 4.0x版以XML 1.0語法重新制定的標籤語法。
- XHTML仍然是使用HTML 4.0版標籤，它就是W3C所制定用來取代HTML 4.0版的下一個世代HTML。撰寫規則如下所示：
 - 不可以省略結尾標籤。
 - 標籤和屬性都是使用小寫的英文字，例如：`<p>`、`<body>`等。
 - 屬性值需要使用引號括起來。
 - name屬性使用id屬性取代。

50

XHTML文件的語法(Ch3-7.xhtml)

```
<?xml version="1.0" encoding="Big5"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0
  Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>這是一個測試網頁</title>
  </head>
  <body>
    <p>歡迎進入我的網頁</p>
  </body>
</html>
```

51

XHTML文件的DTD

Strict DTD

- XHTML使用HTML標籤和CSS，請使用 Strict DTD，此時的宣告如下所示：

```
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

52

XHTML文件的DTD

Transitional DTD

- 瀏覽程式如果不支援CSS，請使用 Transitional DTD，此時的宣告如下所示：

```
<!DOCTYPE html
PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-
transitional.dtd">
```

53

XHTML文件的DTD

Frameset DTD

- XHTML如果使用框架頁，請使用Frameset DTD，此時的宣告如下所示：

```
<!DOCTYPE html
PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">
```

54

XHTML就是良好格式的HTML

- 標籤和屬性名稱都需要小寫
- XHTML文件需要根元素
- 標籤需要結尾標籤
 - 正確：
`<p>TEST</p>`
 - 錯誤：
`<p>TEST`
- 巢狀標籤不能重疊
 - 正確：
`ABC<i>DEF</i>GHI`
 - 錯誤：
`ABC<I>DEFGHI</I>`

55

XHTML就是良好格式的HTML

- 屬性值需要使用引號
- 屬性一定需要擁有屬性值
 - `<option selected= "selected" >`
 - `<input type= "radio" checked= "checked" >`
- Script程式碼需要使用CDATA
 - `<script>`
 - `<![CDATA[`
 - `.....`
 - `]]>`
 - `</script>`

56

習題

- (1) 請為下列資料建立一份XML文件
 - (A) CD 曲目 (可參考下列網址的曲目：
<http://www.books.com.tw/exep/cdfile.php?item=0020115927>)
本題請存檔為(學號_ch3_1_1.xml)
 - (B) 課本目錄
本題請存檔為(學號_ch3_1_2.xml)
- 利用屬性代替元素改寫上列兩個XML文件
本題請存檔為(學號_ch3_2_1.xml及學號_ch3_2_2.xml)