

XML Data

DTDs,
IDs & IDREFs

“Well-Formed” XML

Adheres to basic structural requirements

- Single root element ✓
- Matched tags, proper nesting ✓
- Unique attributes within elements ✓

```
<?xml version="1.0" ?>
<!-- Bookstore with no DTD -->
- <Bookstore>
- <Book ISBN="ISBN-0-13-713526-2" Price="85" Edition="3rd">
  <Title>A First Course in Database Systems</Title>
  - <Authors>
    - <Author>
      <First_Name>Jeffrey</First_Name>
      <Last_Name>Ullman</Last_Name>
    </Author>
    - <Author>
      <First_Name>Jennifer</First_Name>
      <Last_Name>Widom</Last_Name>
    </Author>
  </Authors>
</Book>
- <Book ISBN="ISBN-0-13-815504-6" Price="100">
  <Remark>Buy this book bundled with "A First Course" -- a great deal!</Remark>
  <Title>Database Systems: The Complete Book</Title>
  - <Authors>
```

“Valid” XML

Adheres to basic structural requirements

➤ Also adheres to content-specific specification

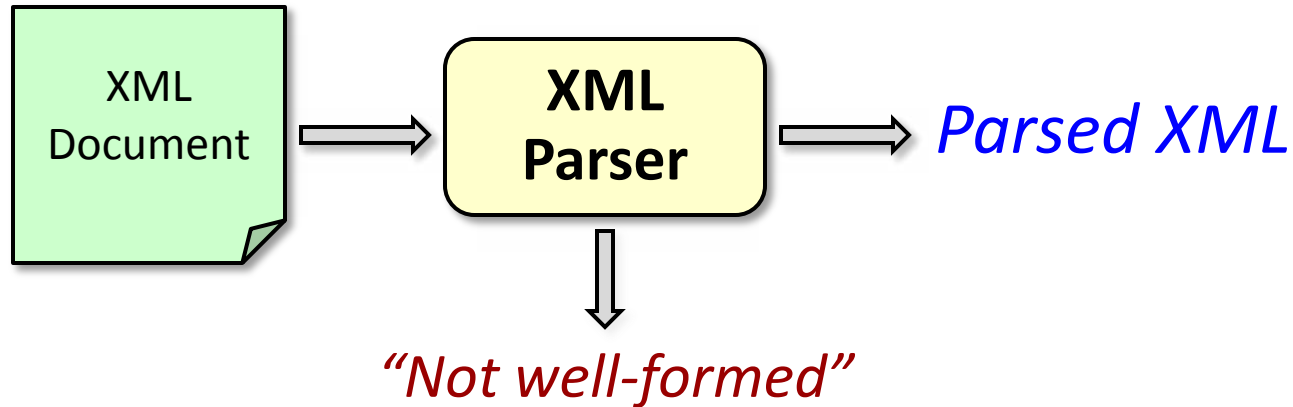
- *Document Type Descriptor* (DTD)
- *XML Schema* (XSD)

```
<?xml version="1.0" ?>
<!-- Bookstore with no DTD -->
- <Bookstore>
- <Book ISBN="ISBN-0-13-713526-2" Price="85" Edition="3rd">
  <Title>A First Course in Database Systems</Title>
  - <Authors>
    - <Author>
      <First_Name>Jeffrey</First_Name>
      <Last_Name>Ullman</Last_Name>
    </Author>
    - <Author>
      <First_Name>Jennifer</First_Name>
      <Last_Name>Widom</Last_Name>
    </Author>
  </Authors>
</Book>
- <Book ISBN="ISBN-0-13-815504-6" Price="100">
  <Remark>Buy this book bundled with "A First Course" -- a great deal!</Remark>
  <Title>Database Systems: The Complete Book</Title>
  - <Authors>
```

“Valid” XML

Adheres to basic structural requirements

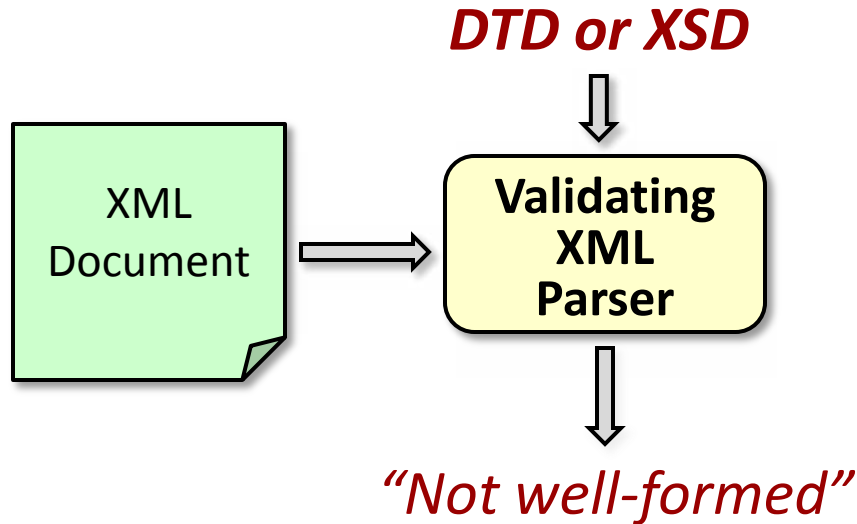
➤ Also adheres to content-specific specification



“Valid” XML

Adheres to basic structural requirements

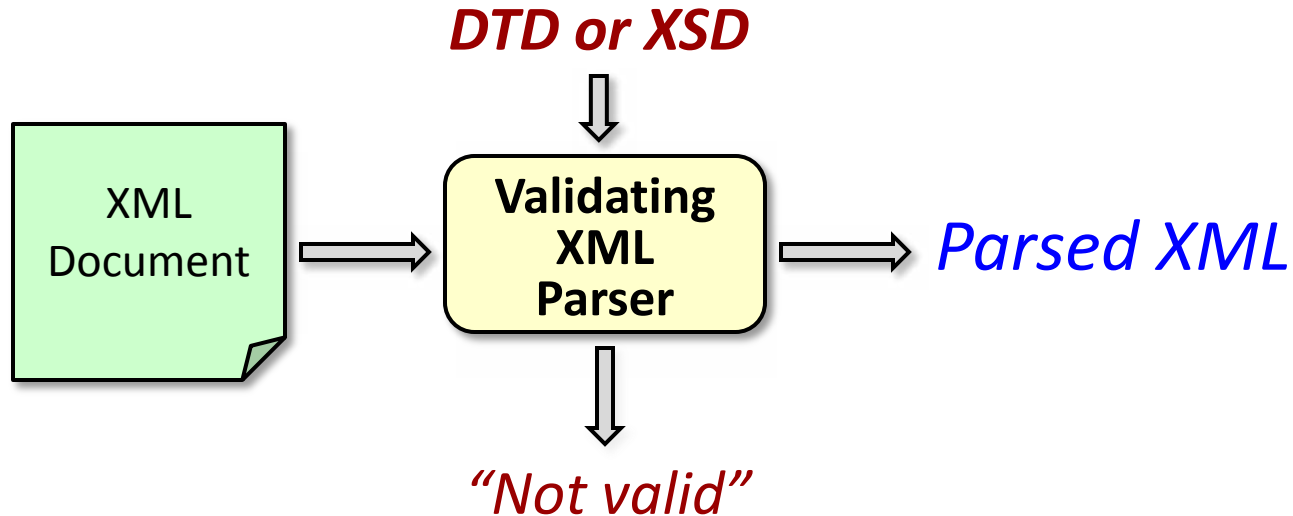
➤ Also adheres to content-specific specification



“Valid” XML

Adheres to basic structural requirements

➤ Also adheres to content-specific specification



Document Type Descriptor (DTD)

- Grammar-like language for specifying elements, attributes, nesting, ordering, #occurrences
- Also special attribute types ID and IDREF(S)
pointers

```
<!DOCTYPE Bookstore [  
  <!ELEMENT Bookstore (Book*, Author*)>  
  <!ELEMENT Book (Title, Remark?)>  
  <!ATTLIST Book ISBN ID #REQUIRED  
    Price CDATA #REQUIRED  
    Authors IDREFS #REQUIRED>  
  <!ELEMENT Title (#PCDATA)>  
  <!ELEMENT Remark (#PCDATA | BookRef)*>  
  <!ELEMENT BookRef EMPTY>  
  <!ATTLIST BookRef book IDREF #REQUIRED>  
  <!ELEMENT Author (First_Name, Last_Name)>  
  <!ATTLIST Author Ident ID #REQUIRED>  
  <!ELEMENT First_Name (#PCDATA)>  
  <!ELEMENT Last_Name (#PCDATA)>  
>
```

DTD/XSD versus none (well-formed)

DTDs, IDs & IDREFs

+ DTD/XSD

Programs can assume structure
"CSS/XSL"
Specification - data exchange
Documentation

Benefits of
"typing"

- DTD/XSD

Flexibility, ease of change
DTDs can be messy - irregular
XSDs
↑
really

Benefits of
"no typing"