

XML Data

Introduction,
Well-formed XML

Extensible Markup Language (XML)

- Standard for data representation and exchange
- Document format similar to HTML
 - Tags describe content instead of formatting
- Also streaming format

```
<?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>
```

Well-Formed XML

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

Annotations:



- Handwritten circled numbers:
  - (1) Circles the entire document structure.
  - (2) Circles the first book entry, with a pink arrow pointing to the handwritten note "No remark".
  - (3) Circles the second book entry, with a pink arrow pointing to the handwritten note "No edition".
- Handwritten notes:
  - "No remark" next to the circled book entry (2).
  - "No edition" next to the circled book entry (3).

```

Basic constructs

- Tagged elements (nested)
- Attributes
- Text

Relational Model versus XML

Well-Formed XML

	Relational	XML
Structure	Tables	Hierarchical Tree, graph
Schema	Fixed in advance	Flexible "Self-describing"
Queries	Simple, nice laugs. 	Less so. 
Ordering	None.	Implied.
Implementation	Native.	Add-on.

“Well-Formed” XML

Adheres to basic structural requirements

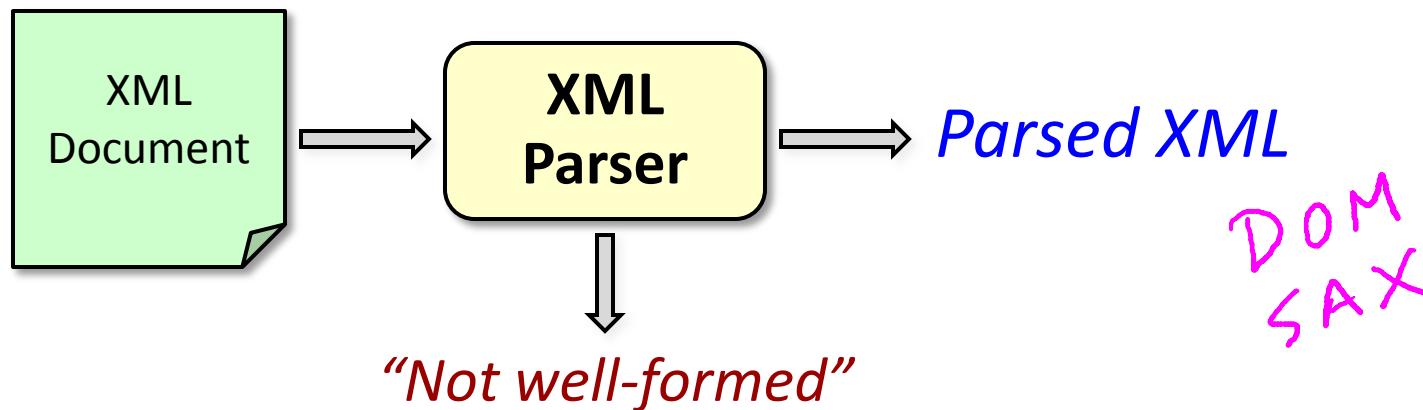
- Single root element
- Matched tags, proper nesting
- Unique attributes within elements

```
<?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>
```

“Well-Formed” XML

Adheres to basic structural requirements

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



Displaying XML

Use rule-based language to translate to HTML

- *Cascading stylesheets* (CSS)
- *Extensible stylesheet language* (XSL)

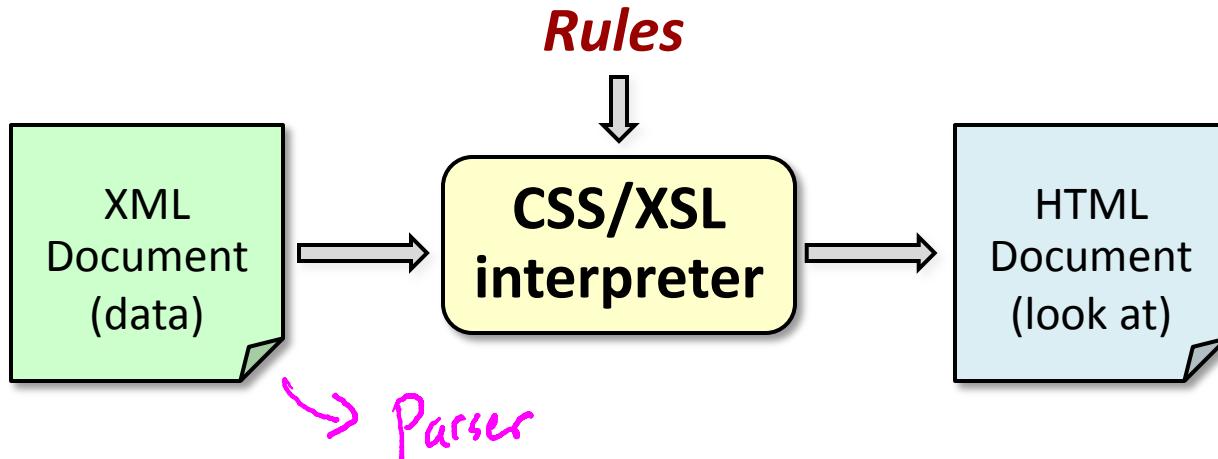


```
<?xml version="1.0" ?>
<!-- Bookstore with no DTD -->
- <Bookstore>
- <Book ISBN="ISBN-0-13-713526-2" Price="85" Edition="3rd">
  <Title>A First Course in Database Systems</Title>
  - <Authors>
    - <Author>
      <First_Name>Jeffrey</First_Name>
      <Last_Name>Ullman</Last_Name>
    </Author>
    - <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>
```

Displaying XML

Use rule-based language to translate to HTML

- *Cascading stylesheets* (CSS)
- *Extensible stylesheet language* (XSL)



Extensible Markup Language (XML)

- Standard for data representation and exchange
- Formal specification is enormous; we cover most important components

```
<?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>
```