

# 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>
```

```
<?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>
      <Author>
        <First_Name>Hector</First_Name>
        <Last_Name>Garcia-Molina</Last_Name>
      </Author>
      <Author>
        <First_Name>Jeffrey</First_Name>
        <Last_Name>Ullman</Last_Name>
      </Author>
      <Author>
        <First_Name>Jennifer</First_Name>

```

Handwritten annotations on the XML code:

- Arrows point to the opening tags of `<Bookstore>`, `<Book>`, `<Author>`, `<First_Name>`, and `<Last_Name>` for the first book.
- A pink oval highlights the `Edition="3rd"` attribute of the first book, with a handwritten note "No remark" and a circled "2" pointing to it.
- A pink oval highlights the `<Remark>` element of the second book, with a handwritten note "No edition" and a circled "3" pointing to it.
- A pink oval highlights the `<Remark>` element of the second book, with a handwritten note "No edition" and a circled "3" pointing to it.

## Basic constructs

- Tagged elements (nested)
- Attributes
- Text

# Relational Model versus XML

Well-Formed XML

	Relational	XML
Structure	Tables	Hierarchical <u>Tree</u> , graph
Schema	Fixed in advance	Flexible "Self-describing"
Queries	Simple, nice langs. 😊	Less so. 😐
Ordering	None.	Implied.
Implementation	Native.	Add-on.

# “Well-Formed” XML

## Adheres to basic structural requirements

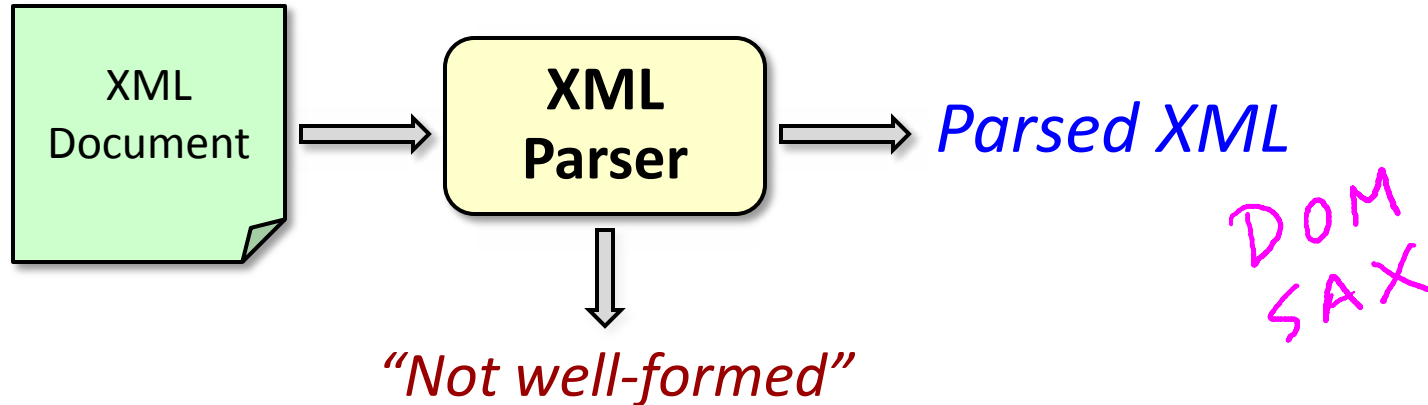
- Single root element
- Matched tags, proper nesting
- Unique attributes within elements

```
<?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)*

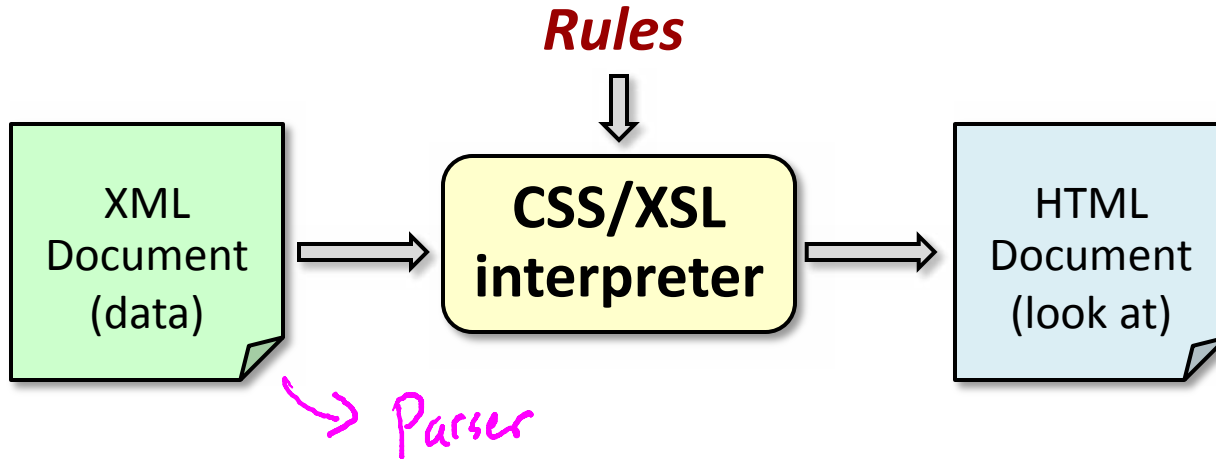


```
<?xml version="1.0" ?>
<!-- Bookstore with no DTD -->
- <Bookstore>
-   <Book ISBN="ISBN-0-13-713526-2" Price="85" Edition="3rd">
       <Title>A First Course in Database Systems</Title>
       <Authors>
       - <Author>
           <First_Name>Jeffrey</First_Name>
           <Last_Name>Ullman</Last_Name>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