

# Знакомство с XML

**Язык XML (*Extensible Markup Language*)** был разработан рабочей группой *XML Working Group* консорциума *World Wide Web Consortium (W3C)*. Вот как описывают его создатели:

*"Расширяемый язык разметки Extensible Markup Language (XML) представляет собой составную часть языка SGML... Он предназначен для облегчения использования языка SGML в Web и выполнения задач, которые в настоящее время реализуются с помощью языка HTML. XML разработан с целью усовершенствовать применение и взаимодействие языков SGML и HTML."*

Описание на языке *XML* представляет собой **операторы**, написанные с соблюдением определенного синтаксиса. Когда вы создаете *XML*-документ, то вместо использования ограниченного набора определенных элементов вы имеете возможность создавать ваши собственные элементы и присваивать им любые имена по вашему выбору – именно поэтому язык *XML* является расширяемым (*extensible*). Следовательно, вы можете использовать *XML* для описания практически любого документа, от музыкальной партитуры до *базы данных*.

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

```
<INVENTORY>
```

```
<BOOK><TITLE>The Adventures of Huckleberry Finn</TITLE>
```

```
  <AUTHOR>Mark Twain</AUTHOR> <BINDING>mass market  
  paperback</BINDING>
```

```
  <PAGES>298</PAGES>
```

```
  <PRICE>$5.49</PRICE>
```

```
</BOOK>
```

```
<BOOK> <TITLE>Moby-Dick</TITLE> <AUTHOR>Herman
```

```
  Melville</AUTHOR> <BINDING>trade paperback</BINDING>
```

```
  <PAGES>605</PAGES> <PRICE>$4.95</PRICE>
```

```
</BOOK>
```

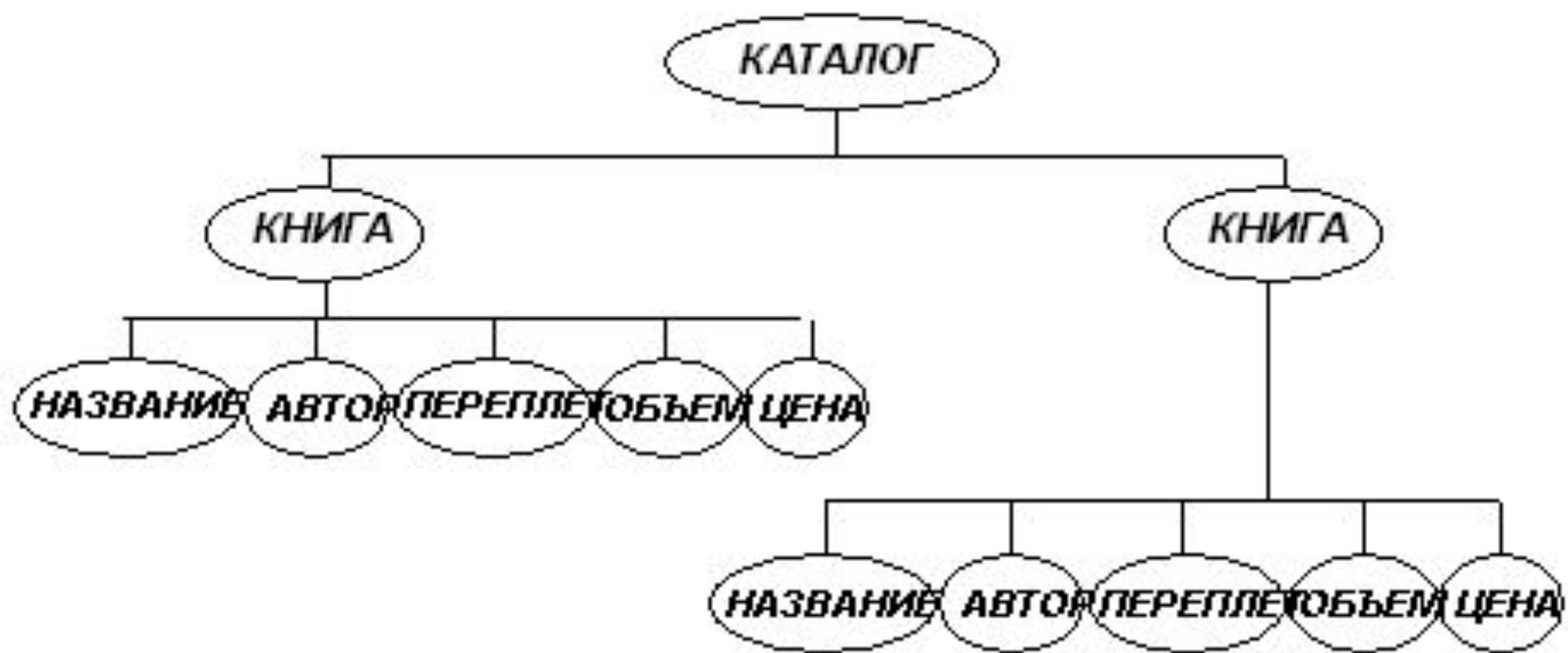
```
<BOOK> <TITLE>The Scarlet Letter</TITLE> <AUTHOR>Nathaniel
```

```
  Hawthorne</AUTHOR> <BINDING>trade paperback</BINDING>
```

```
  <PAGES>253</PAGES> <PRICE>$4.25</PRICE>
```

```
</BOOK>
```

```
</INVENTORY>
```



# Официальные концептуальные цели XML

- XML должен стать языком прямого использования в Internet.
- XML будет поддерживать большое количество приложений.
- XML будет совместим с SGML.
- Будет легче писать программы, обрабатывающие XML-документы.
- Количество дополнительных функций в XML должно быть минимальным, а в идеале – нулевым.

# Официальные концептуальные цели XML

- XML-документы должны быть понятными и ясными для пользователя.
- Разработка XML должна быть завершена достаточно быстро.
- Язык XML должен быть формальным и кратким.
- XML-документы будет проще создавать.
- Сжатой форме при XML-разметке придается минимальное значение.

# Стандартные XML-приложения

Индивидуальный пользователь, компания или комитет по стандартам может определить необходимый набор элементов XML и структуру документа, которые будут применяться для особого класса документов. Подобный набор элементов и описание структуры документа называют XML-приложением или XML-словарем.

# Стандартные XML-приложения

XML-приложение обычно определяется созданием описателя типа документа (DTD), который является допустимым компонентом XML-документа. DTD построен по схеме базы данных: он устанавливает и определяет имена элементов, которые могут быть использованы в документе, порядок, в котором элементы могут появляться, доступные к применению атрибуты элементов и другие особенности документа.

# XML-приложения, повышающие качество XML-документов

- **Extensible Stylesheet Language (XSL)** позволяет вам создавать мощные стилевые таблицы с использованием синтаксиса XML.
- **XML Schema** позволяет разрабатывать подробные схемы для ваших XML-документов с использованием стандартного синтаксиса XML, что является более мощной альтернативой применения DTD.
- **XML Linking Language (XLink)** дает возможность связывать ваши XML-документы. Он поддерживает множественные целевые ссылки и другие полезные функции, обеспечивая большую свободу по сравнению с механизмом организации ссылок в HTML.
- **XML Pointer Language (XPointer)** позволяет определять гибкие целевые ссылки. При совместном использовании XPointer и XLink вы можете организовывать ссылки на любое место в целевом документе – а не только переходы к специально выделенным пунктам.

# Реальное использование XML

- Работа с базами данных.
- Структурирование документов.
- Работа с векторной графикой
- Мультимедиа презентации
- Описание каналов.
- Описание программных пакетов и их взаимосвязей.
- Взаимодействие приложений через Web с использованием XML-сообщений.
- Отправка электронных бизнес-карт через e-mail.
- Обмен финансовой информацией.

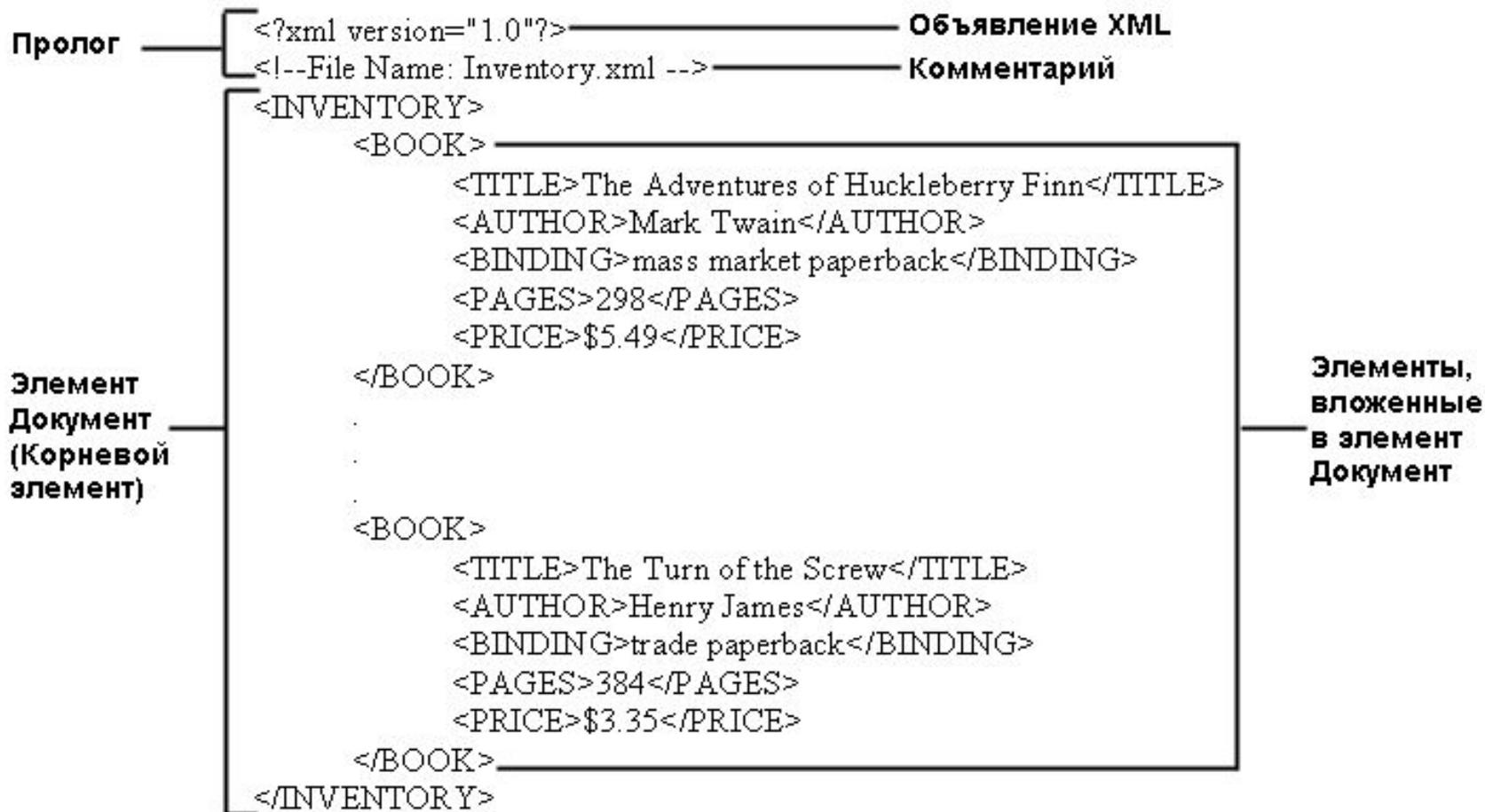
# Реальное использование XML

- Создание, управление и использование сложных цифровых форм для коммерческих Internet-транзакций.
- Обмен запросами по приему на работу и резюме
- Форматирование математических формул и научной информации в Web.
- Описание молекулярных структур
- Кодирование и отображение информации о ДНК, РНК и цепочках
- Кодирование генеалогических данных
- Обмен астрономическими данными
- Создание музыкальных партитур

# Реальное использование XML

- Работа с голосовыми сценариями для доставки информации по телефону.
- Обработка и доставка информации курьерскими службами.
- Представление рекламы в прессе в цифровом формате.
- Заполнение юридических документов и электронный обмен юридической информацией (XCL – XML Court Interface).
- Кодирование прогнозов погоды.
- Обмен страховой информацией.
- Обмен новостями и информацией с использованием открытых Web-стандартов (XMLNews).

# Анатомия XML-документа



# Пролог

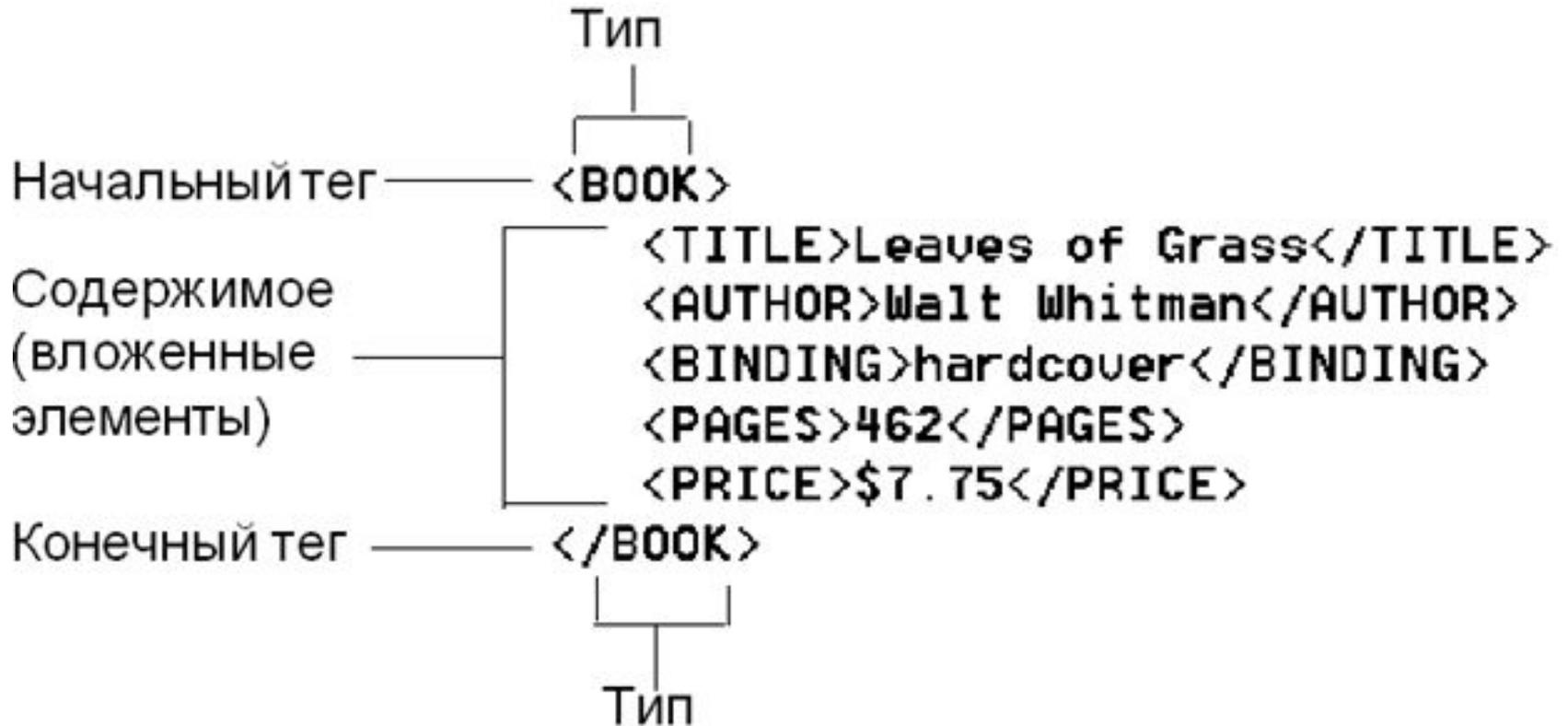
- Первая строка представляет собой объявление *XML*, указывающее на то, что это *XML*-документ и содержащее номер версии.
- Вторая строка пролога состоит из пробела.
- Третья строка пролога представляет собой комментарий.
- *объявление типа документа*
- одна или несколько инструкций по обработке, содержащих информацию о порядке проходов при обработке приложения XML-процессором.

# Элемент Документ

Второй основной частью XML-документа является единый элемент Документ, или корневой элемент, который в свою очередь содержит дополнительные элементы.

В XML-документе элементы определяют его логическую структуру и несут в себе информацию, содержащуюся в документе (в нашем примере это *информация* о книгах, такая как название, *автор*, цена). Типовой элемент состоит из начального тега, содержимого элемента и конечного тега. Содержимым элемента могут быть символьные данные, другие (вложенные) элементы, либо сочетание данных и вложенных элементов.

# Элемент Документ



# Элемент Документ



# Некоторые базовые правила XML

- Документ должен иметь только один элемент верхнего уровня (элемент Документ или корневой элемент).
- Элементы должны быть вложены упорядоченным образом.
- Каждый элемент должен иметь начальный и конечный тег.
- Имя типа элемента в начальном теге должно в точности соответствовать имени в соответствующем конечном теге.
- Имена типов элементов чувствительны к регистру, в котором они набраны.