

Описание формата FB2 от Sclex

Материал из FictionBook

FB2-файл - это файл XML

FB2-документ является файлом в формате XML.

XML (англ. Extensible Markup Language ? расширяемый язык разметки; произносится [икс-эм-э?л]) ? рекомендованный Консорциумом Всемирной паутины язык разметки, фактически представляющий из себя свод общих синтаксических правил. XML предназначен для хранения структурированных данных, для обмена информацией между программами, а также для создания на его основе более специализированных языков разметки (например, XHTML), иногда называемых словарями.

Ниже приведён пример простого кулинарного рецепта, размеченного с помощью XML:

```
<?xml version="1.0" encoding="UTF-8"?>
<Recipe name="хлеб" prep_time="5 мин" cook_time="3 час">
  <название>Простой хлеб</название>
  <ingredient amount="3" unit="стакан">Мука</ingredient>
  <ingredient amount="0.25" unit="грамм">Дрожжи</ingredient>
  <ingredient amount="1.5" unit="стакан">Тёплая вода</ingredient>
  <ingredient amount="1" unit="чайная ложка">Соль</ingredient>
  <Instructions>
    <step>Смешать все ингредиенты и тщательно замесить.</step>
    <step>Закрыть тканью и оставить на один час в тёплом помещении.</step>
    <step>Замесить ещё раз, положить на противень и поставить в духовку.</step>
  </Instructions>
</Recipe>
```

Обратите внимание, что названия и значения элементов и атрибутов могут состоять не только из букв латинского алфавита, но десятичным разделителем может быть только точка.

Первая строка XML-документа называется **объявлением XML** ? это необязательная строка, указывающая версию стандарта XML (обычно это 1.0), также здесь может быть указана кодировка символов и внешние зависимости.

Остальная часть этого XML-документа состоит из вложенных *элементов*, некоторые из которых имеют *атрибуты* и *содержимое*. **Элемент** обычно состоит из открывающего и закрывающего тегов (меток), обрамляющих текст и другие элементы. **Открывающий тег** состоит из имени элемента в угловых скобках, например ?<step>; **закрывающий тег** состоит из того же имени в угловых скобках, но перед именем ещё добавляется косая черта, например ?</step>?. **Содержимым элемента** называется всё, что расположено между открывающим и закрывающим тегами, включая текст и другие (вложенные) элементы. Ниже приведён пример XML-элемента, который содержит открывающий тег, закрывающий тег и содержимое элемента:

```
<step>Замесить ещё раз, положить на противень и поставить в духовку.</step>
```

Кроме содержания у элемента могут быть **атрибуты** ? пары имя-значение, добавляемые в открывающий тег после названия элемента. Значения атрибутов всегда заключаются в кавычки (одинарные или двойные), одно и то же имя атрибута не может встречаться дважды в одном элементе. Не рекомендуется из соображений читабельности использовать разные типы кавычек для значений атрибутов одного тега.

```
<ingredient amount="3" unit="стакан">Мука</ingredient>
```

В приведённом примере у элемента `?ingredient?` есть два атрибута: `?amount?`, имеющий значение `?3?`, и `?unit?`, имеющий значение `?стакан?`. С точки зрения XML-разметки, приведённые атрибуты не несут никакого смысла, а являются просто набором символов.

Кроме текста элемент может содержать другие элементы:

```
<Instructions>
  <step>Смешать все ингредиенты и тщательно замесить.</step>
  <step>Закрыть тканью и оставить на один час в тёплом помещении.</step>
  <step>Замесить ещё раз, положить на противень и поставить в духовку.</step>
</Instructions>
```

В данном случае элемент `?Instructions?` содержит три элемента `?step?`. XML не допускает перекрывающихся элементов. Например, приведённый ниже фрагмент некорректен, так как элементы `?em?` и `?strong?` перекрываются.

```
<!-- ВНИМАНИЕ! Некорректный XML! -->
<p>Обычный <em>акцентированный <strong>выделенный и акцентированный</em> выделенный</strong></p>
```

Каждый XML-документ должен содержать в точности один **корневой элемент** или *document element*), таким образом, следующий фрагмент не может считаться корректным XML-документом.

```
<!-- ВНИМАНИЕ! Некорректный XML! -->
<thing>Сущность ?1</thing>
<thing>Сущность ?2</thing>
```

Для обозначения элемента без содержания, называемого **пустым элементом**, допускается применять особую форму записи, состоящую из одного тега, в котором после имени элемента ставится косая черта. Следующие фрагменты полностью равнозначны:

```
<foo></foo>
```

```
<foo/>
```

В XML определены два метода записи специальных символов: ссылка на сущность и ссылка по номеру символа. **Сущностью** в XML называются именованные данные, обычно текстовые, в частности спецсимволы. **Ссылка на сущность** указывается в том месте, где должна быть сущность и состоит из амперсанда (`?&?`), имени сущности и точки с запятой (`?;?`). В XML есть несколько предопределённых сущностей, таких как `?lt?` (ссылаться на неё можно написав `?<?`) для левой угловой скобки и `?amp?` (ссылка `?&?`) для амперсанда, возможно также определять собственные сущности. Помимо записи с помощью сущностей отдельных символов, их можно использовать для записи часто встречающихся текстовых блоков. Ниже приведён пример использования предопределённой сущности для избежания использования знака амперсанда в названии:

```
<company-name>AT&amp;T</company-name>
```

Полный список предопределённых сущностей состоит из `&` (`?&?`), `<` (`?<?`), `>` (`?>?`), `'` (`?'??`), и `"` (`?"??`) ? последние две полезны для записи разделителей внутри значений атрибутов. Определить свои сущности можно в DTD-документе. Иногда бывает необходимо определить неразрывный пробел, который в HTML обозначается как ` `; в XML его записывают ` `;

Ссылка по номеру символа выглядит как ссылка на сущность, но вместо имени сущности указывается символ # и число (в десятичной или шестнадцатеричной записи), являющееся номером символа в кодовой таблице Юникод. Это обычно символы, которые невозможно закодировать напрямую, например буква арабского алфавита в ASCII-кодированном документе. Амперсанд может быть представлен следующим образом:

```
<company-name>AT&#038;T</company-name>
```

Существует ещё множество правил, касающихся составления корректного XML-документа, но целью данного краткого обзора было лишь показать основы, необходимые для понимания структуры XML-документа.

Общая структура документа. Кодировки

FB2-документ состоит из объявления XML и корневого элемента `<FictionBook>`. То есть выглядит это так:

```
<?xml version="1.0" encoding="windows-1251"?>
<FictionBook xmlns="http://www.gribuser.ru/xml/fictionbook/2.0"
  xmlns:l="http://www.w3.org/1999/xlink">
  ...
</FictionBook>
```

Как видно, в первой строке в значении атрибута `encoding` надо правильно указать кодировку документа. В приведенном примере это `windows-1251`, а в другом случае могла бы быть и `utf-8`. `utf-8` также является кодировкой по умолчанию, используемой в случае отсутствия атрибута `encoding` или всего XML-объявления.

Ниже указано, какая кодировка каким языкам соответствует:

```
windows-1251 - английский и кириллица (русский, болгарский и т.д.)
windows-1252 - западно-европейские языки (английский, французский, испанский и т.д.)
windows-1250 - английский и восточно-европейские (польский, чешский и т.д.)
```

Раздел FictionBook

Этот раздел начинается сразу за первой строкой и открывается таким образом:

```
<FictionBook xmlns="http://www.gribuser.ru/xml/fictionbook/2.0"
  xmlns:l="http://www.w3.org/1999/xlink">
```

А закрывается в самом конце FB2-документа.

В `xmlns:l` вместо `l` может быть написано и другое имя, главное, чтобы потом оно использовалось в ссылках по всему документу. То есть, если объявлено `xmlns:l="http://www.w3.org/1999/xlink"`, то ссылки надо оформлять как `<a l:href=...>...`, а если объявлено `xmlns:xlink="http://www.w3.org/1999/xlink"`, то надо так: `<a xlink:href=...>...`. Аналогично и для ссылок в тэгах `<image/>`.

Раздел `FictionBook` состоит из вложенных подразделов в указанном ниже порядке:

1. **<description>** - который описывает заголовок документа. *Одно и только одно вхождение.* (фразы вроде "одно и только одно вхождение" говорят, сколько раз подряд может идти данный тэг в данном месте документа)
2. **<body>** - описывает тело документа. *Одно или более вхождений.*
3. **<binary>** - содержит приложенные к файлу двоичные объекты - картинки и прочее. *Любое число вхождений.*

Иными словами, присутствуют как минимум разделы **<description>** с **<body>**, а остальное - по необходимости.

Пример:

```
<?xml version="1.0" encoding="windows-1251"?>
<FictionBook xmlns="http://www.gribuser.ru/xml/fictionbook/2.0"
  xmlns:l="http://www.w3.org/1999/xlink">
  <description>
    ...
  </description>
  <body>
    ...
  </body>
  <body name="notes">
    ...
  </body>
  <binary id="cover.jpg" content-type="image/jpeg"/>9j/
    4AAQSkZJRgABAgAAZABkAAD/
    ...
  </binary>
</FictionBook>
```

Поля раздела description

Раздел **<description>** состоит из перечисленных ниже частей в указанном порядке:

1. **<title-info>** - данные о книге. *Одно и только одно вхождение.*
2. **<src-title-info>** - данные об источнике книги (до перевода). *От нуля до одного вхождений.*
3. **<document-info>** - информация об FB2-документе. *Одно и только одно вхождение.*
4. **<publish-info>** - сведения об издании книги, которая была использована как источник при подготовке документа. *От нуля до одного вхождений.*
5. **<custom-info>** - произвольная информация. *Любое число вхождений.*

То есть обязательны разделы **<title-info>** и **<document-info>**, а остальные добавляются по необходимости.

Рассмотрим пример заполнения его полей (они должны следовать в указанном порядке):

```
<description>
<title-info>
  <genre>sci_medicine</genre>
  <genre>science</genre>
  <author>
    <first-name>Борис</first-name>
    <last-name>Сергеев</last-name>
  </author>
  <book-title>Ум хорошо:</book-title>
  <annotation>
    <p>Об организации высших психических функций.</p>
  </annotation>
  <date value="1984-01-01">1984</date>
  <coverpage><image l:href="#cover.jpg"/></coverpage>
  <lang>ru</lang>
  <src-lang>ru</src-lang>
  <sequence name="Раскрывая тайны психики" number="2"/>
```

```
</title-info>
<document-info>
  <author>
    <nickname>Sclex</nickname>
    <email>sclex_at_inbox.ru</email>
  </author>
  <program-used>Dn/2, Opera 8.50, Bred3</program-used>
  <date value="2006-05-21">2006-05-21</date>
  <id>91824ABE-EEF7-48C3-A719EFC4824A</id>
  <version>1.1</version>
  <history>
    <p>1.0 - подготовка fb2 by Sclex</p>
    <p>1.1 - вычитка ошибок by Sclex</p>
  </history>
</document-info>
<publish-info>
  <book-name>Борис Федорович Сергеев. Ум хорошо:</book-name>
  <publisher>Молодая гвардия</publisher>
  <city>М.</city>
  <year>1984</year>
</publish-info>
</description>
```

Рассмотрим последовательно смысл значений тэгов заголовка. (Некоторые из тэгов отсутствуют в примере, но описаны ниже).

Внутри раздела **<title-info>**:

<genre> - указывает жанр книги. Список возможных значений приведен в соответствующем разделе данного описания. *Одно или более вхождений*, то есть книга может относиться к нескольким жанрам одновременно. В нашем примере это `sci_medicine` и `science`, то есть книга относится к жанрам "Медицина" и "Научно-образовательная: Прочее".

<author> - содержит информацию об авторе книги. Возможны такие вложенные тэги как **<first-name>** - имя, **<last-name>** - фамилия, **<middle-name>** - отчество, **<nickname>** - ник, и **<email>** - адрес электронной почты. *Одно или более вхождений*, так как может быть несколько авторов у одной книги. В нашем примере определены имя Борис и фамилия Сергеев.

<book-title> - описывает название книги. *Одно и только одно вхождение*. В нашем примере книга называется "Ум хорошо:".

<annotation> - аннотация книги. *Возможны в любом количестве и порядке* следующие вложенные тэги: **<p>**, **<poem>**, **<cite>**, **<subtitle>**, **<table>**, **<empty-line/>** с соответствующим содержанием. *От нуля до одного вхождения*.

<keywords> - содержит ключевые слова к данной книге для поисковых систем. Слова должны быть в виде неформатированного текста. *От нуля до одного вхождения*.

<date> - хранит дату написания книги. В свойстве **value** дата описывается в формате ГГГГ-ММ-ДД, а внутри тэга дата задается в текстовом виде. *От нуля до одного вхождения*. В нашем примере свойство **value** имеет значение 1984-01-01, то есть первое января 1984-го года, а внутри тэга прописано значение 1984.

<coverpage> - обложка книги. Внутри может содержать только тэг **<image/>**. *От нуля до одного вхождения*. В нашем примере ссылка указывает на картинку `cover.jpg`, приложенную к документу (значок # говорит, что эта ссылка локальная, то есть адресует в пределах документа).

<lang> - сообщает язык книги в документе, то есть язык после перевода. *Одно и только одно вхождение*. В нашем примере это **ru**, то есть русский язык. Полный список кодов приведен в соответствующем разделе данного описания.

<src-lang> - определяет язык, на котором исходно написана книга, то есть язык до перевода. *От нуля до одного вхождений.* Те же коды, что для тэга **<lang>**. В нашем примере это **ru**, то есть русский язык. Английскому языку соответствует код **en**.

<translator> - информация о переводчике книги. Возможны такие вложенные тэги как **<first-name>** - имя, **<last-name>** - фамилия, **<middle-name>** - отчество, **<nickname>** - ник, и **<email>** - адрес электронной почты. *Любое число вхождений.*

<sequence> - сведения о том, к каким сериям относится книга. Имеет такие свойства, как **name** (текстовая строка) - название серии, и **number** (целое число) - порядковый номер книги в серии. *Любое число вхождений.* В нашем примере описана серия "Раскрывая тайны психики", и указан номер книги в серии - вторая.

Внутри раздела **<document-info>**:

<author> - содержит информацию об авторе документа. Возможны такие вложенные тэги как **<first-name>** - имя, **<last-name>** - фамилия, **<middle-name>** - отчество, **<nickname>** - ник, и **<email>** - адрес электронной почты. *Одно или более вхождений.* В нашем примере определены ник Sclex и адрес электронной почты sclex_at_inbox.ru.

<program-used> - в виде простого перечисляет программы, использованные при создании FB2-документа. *От нуля до одного вхождений.* В нашем примере это Dn/2, Opera 8.50 и Bred3.

<date> - хранит дату создания документа. В свойстве **value** дата описывается в формате ГГГГ-ММ-ДД, а внутри тэга дата задается в текстовом виде. *Одно и только одно вхождение.* В нашем примере свойство **value** имеет значение 2006-05-21, то есть 21-е мая 2006-го года, и такое же значение прописано внутри тэга.

<src-url> - URL страницы, откуда взят текст для подготовки документа. *Любое число вхождений.*

<src-ocr> - автор текста, который был использован при подготовке документа. Но не тот автор, который написал книгу, а тот, который сканировал ее и подготовил электронный текст. *От нуля до одного вхождений.*

<id> - уникальный идентификатор документа FB2. Каждый отдельный FB2-документ должен иметь собственный ID, отличающийся от ID других документов. Это значит, что при изменении книги, которая есть в библиотеке, ID нужно сохранить. Если же создается новая книга, это поле необходимо заполнить уникальной записью. Формально разрешено использовать внутри тэга ID любые символы в любом количестве, но рекомендуется употреблять символы латиницы, цифры, и знаки подчеркивание и минус ("_", "-"). Других требований к идентификатору нет. Как вариант, можно заполнить поле согласно времени создания документа по формату YYYYMMDDhhmmss, например 20060728182041. *Одно и только одно вхождение.*

<version> - версия документа в текстовом виде. Описывается дробным числом. *Одно и только одно вхождение.* В нашем случае - 1.1.

<history> - история создания и изменения документа. Содержит форматированный текст. Возможны в любом количестве и порядке следующие вложенные тэги: **<p>**, **<poem>**, **<cite>**, **<subtitle>**, **<table>**, **<empty-line/>** с соответствующим содержанием. *От нуля до одного вхождений.*

Внутри тэга **<publish-info>**:

<book-name> - название оригинальной (бумажной) книги. *От нуля до одного вхождений.* В примере это "Борис Федорович Сергеев. Ум хорошо:".

<publisher> - название издательства. *От нуля до одного вхождений.* В примере - Молодая гвардия.

<city> - город, в котором издана книга. *От нуля до одного вхождений*. В нашем случае это М., то есть Москва.

<year> - год издания книги. *От нуля до одного вхождений*. В нашем случае - 1984.

Тело документа (структура разделов **body** и **section**)

Тэг <body> предназначен для описания собственно тела документа. Этот тэг может иметь свойство **name**, которое описывает имя раздела. Например:

```
<body name="notes">
```

Внутри раздела <body> могут идти в указанном порядке:

1. Картинка <image/>, которая будет отображаться вверху раздела. *От нуля до одного вхождений*.
2. Заголовок <title>. *От нуля до одного вхождений*.
3. Эпиграф <epigraph>. *Любое число вхождений*.
4. Раздел <section>. *Одно или более вхождений*.

Откуда видно, что обязателен только раздел <section>.

Для раздела <section> имя описывается свойством **id**. Например так:

```
<section id="note33">
```

Раздел <section> состоит из следующих частей в указанном порядке:

1. Заголовок <title>. *От нуля до одного вхождений*.
2. Эпиграф <epigraph>. *Любое число вхождений*.
3. Картинка <image/>, которая будет отображаться вверху раздела. *От нуля до одного вхождений*.
4. Аннотация <annotation>, описывающая раздел. *От нуля до одного вхождений*.
5. Одно из двух. Либо подраздел <section> (*одно или более вхождений*); либо форматированный текст с картинками - тэги <p>, <image>, <poem>, <subtitle>, <cite>, <empty-line/>, <table> (*одно или более вхождений*), но первый тэг не может быть <image>.

Следовательно, обязательно наличие либо вложенного подраздела <section>, либо форматированного текста с картинками.

Разделов <body> в документе может быть несколько. Это используется, например, для оформления примечаний (footnotes). По умолчанию пользователю показывается первый раздел <body>, а содержание остальных должно быть доступно по гиперссылкам. Имя раздела должно отражать смысл его содержания, но для основного раздела это необязательно.

Разделы <section> должны иметь вложенную структуру согласно иерархии оглавления. Например, если документ делится на части, которые в свою очередь делятся на главы, то части будут описаны в разделе <section> первого уровня, а главы - в подразделе второго уровня. Проще понять это на примере:

```
<body>
  <title><p>Борис Сергеев</p><p>Парадоксы мозга</p></title>
  <section>
    <title><p>Часть 1. Великие невежды.</p></title>
    <section>
      <title><p>Глава 1. Во тьме веков.</p></title>
      <p>Когда люди стали о себе задумываться? ...</p>
      ...
    </section>
  </section>
</body>
```

Принято названия глав писать маленькими буквами, начиная с большой.

Еще принято кроме того, что указано в заголовке, давать имя автора и название книги в тексте документа до того, как открывается **<section>** первого уровня, то есть в **<title>**, относящемся к **<body>**. Это видно из вышеприведенного примера.

Форматированный текст

Форматированный текст представляет собой последовательность фрагментов, ограниченных тэгами **<p>**, **<poem>**, **<subtitle>**, **<cite>**, **<table>**. Также могут входить пустые строки - **<empty-line/>**.

Тэг **<p>** предназначен для описания одного абзаца текста книги.

Например:

```
<p>Вот абзац текста.</p>
<p>Еще один абзац текста.</p>
```

Тэг **<p>** может быть вложен в тэги **<title>**, **<cite>**, **<epigraph>**, **<annotation>**, **<section>**.

Пустые строки

Если нужно в форматированном тексте вставить пустую строку, это делается с помощью тэга **<empty-line/>**. Один такой тэг - одна пустая строка. Закрывающий тэг в данном случае не требуется.

Например:

```
<p>Абзац до пустой строки.</p>
<empty-line/>
<p>А вот это уже абзац после пустой строки.</p>
```

Жирный и курсивный текст

Для форматированного текста, а именно внутри тэгов **<p>**, **<v>**, **<subtitle>**, **<text-author>** возможно описать текст как жирный или курсивный. Жирный текст следует выделить тэгом ****, а курсивный - тэгом **<emphasis>**.

Например:

```
<p>Вот это слово выведем <strong>жирным</strong>,
а вот это - <emphasis>курсивным</emphasis>.</p>
```

Верхние и нижние индексы

В форматированном тексте (внутри тэгов `<p>`, `<v>`, `<subtitle>`, `<text-author>`) можно задать для текста начертание как верхний или нижний индекс. Для нижних индексов (уменьшенные буквы внизу строки) используется выделение тэгом `<sub>`, а для верхних индексов (уменьшенные буквы вверху строки) - выделение тэгом `<sup>`.

Пример:

```
<p>2<sup>4</sup>=16</p>
<p>a<sub>1</sub>+a<sub>2</sub>=b</p>
```

Перечеркнутый текст

В форматированном тексте (внутри тэгов `<p>`, `<v>`, `<subtitle>`, `<text-author>`) можно задать для текста отображение перечеркнутым шрифтом. Для этого используется выделение тэгом `<strikethrough>`.

```
<p><strikethrough>Перечеркнутый текст.</strikethrough></p>
```

Программный код

В форматированном тексте (внутри тэгов `<p>`, `<v>`, `<subtitle>`, `<text-author>`) можно логически разметить текст, который при просмотре файла должен быть отображен как программный код. Для этого используется выделение тэгом `<code>`. Читалкам рекомендуется текст внутри этого тэга оформлять при выводе моноширинным шрифтом (то есть шрифтом, символы которого имеют одинаковую ширину).

```
<p><code>Assign(f, 'name.txt');</code></p>
```

Заголовки

Заголовки частей книги описываются тэгом `<title>`, который содержит в себе форматированный текст. Тэг `<title>` как заголовок части книги может быть вложенным в тэги `<body>` и `<section>`, причем *более одного заголовка в разделе не допускается, а может его и не быть вовсе*.

В тэг `<title>` могут быть вложены только тэги `<p>` и/или `<empty-line>`.

Вот пример правильного оформления заголовка с пустой строкой:

```
<title>
<p>Глава 1</p>
<empty-line/>
<p>От Александрии до Лондона</p>
</title>
```

Тэг `<title>` может быть вложен в тэги `<body>`, `<section>`, `<poem>`, `<stanza>`.

Подзаголовки

Подзаголовки описываются тэгом `<subtitle>`, который содержит простой (неформатированный) текст. Тэг `<subtitle>` может быть вложенным в тэги `<body>`, `<section>`, `<cite>` и `<stanza>`.

Обычно описание подзаголовка выглядит так:

```
<subtitle>* * *<subtitle>
```

Подзаголовком, в отличие от нормального заголовка, оформляется элемент оглавления, незначимый для структуры документа. Обычно это те самые три звездочки или х х х.

Стихи

Чтобы разместить стихотворение в тексте документа, используют тэг **<poem>**. Он может нести внутри себя следующие части в указанном порядке:

1. Заголовок **<title>**. *От нуля до одного вхождений.*
2. Эпиграф **<epigraph>**. *Любое число вхождений.*
3. Строфы **<stanza>**. *Одно или более вхождений.* Строфа **<stanza>** - это группа строк стихотворения, отделенная пустым промежутком от остальных строк. Отдельная строка стихотворения помечается тэгом **<v>**, который должен быть вложен в **<stanza>**. Кроме того в состав **<stanza>** могут входить тэги **<title>** (заголовок) и **<subtitle>** (подзаголовок).
4. Автор стихотворения **<text-author>**. *Любое число вхождений.*
5. Дата написания **<date>**. *От нуля до одного вхождений.*

То есть обязательно присутствует как минимум одна строфа **<stanza>**, а остальное добавляется при необходимости.

Например, стихотворение из двух строф:

*Наша Таня громко плачет,
Уронила в речку мячик.*

*Тише, Танечка, не плачь,
Не утонет в речке мяч.*

Следует разметить таким образом:

```
<poem>
  <stanza>
    <v>Наша Таня громко плачет,</v>
    <v>Уронила в речку мячик.</v>
  </stanza>
  <stanza>
    <v>Тише, Танечка, не плачь,</v>
    <v>Не утонет в речке мяч.</v>
  </stanza>
</poem>
```

Эпиграфы

Для оформления эпиграфов предусмотрен тэг **<epigraph>**, который должен содержать внутри себя форматированный текст - тэги **<p>**, **<poem>**, **<cite>**, **<empty-line>** (*любое число вхождений*), а затем автора текста **<text-author>** (*тоже любое число вхождений*).

Пример:

```
<epigraph>
  <p>Читателям, которые станут доискиваться, какие реальные
  люди и места описаны здесь под вымышленными именами и названиями,
  я бы посоветовал посмотреть вокруг себя и заглянуть в собственную душу,
  так как в этом романе рассказано о том, что происходит сегодня почти
  во всей Америке.</p>
</epigraph>
```

Тэг **<epigraph>** может быть вложен в тэги **<body>**, **<section>** и **<poem>**.

Цитированный текст

Для разметки цитированного текста существует тэг **<cite>**, который может содержать форматированный текст - вложенные тэги **<p>**, **<poem>**, **<empty-line/>**, **<subtitle>**, **<table>** (любое число вхождений), а затем автора цитаты **<text-author>** (любое число вхождений).

Пример:

```
<cite>
  <p>Время - деньги.</p>
  <text-author>Бенджамин Франклин</text-author>
</cite>
```

Автор цитаты

Внутри тэгов **<poem>**, **<cite>** и **<epigraph>** возможно указать автора соответственно стихотворения, цитаты или эпиграфа. Для этого служит тэг **<text-author>**. Внутри **cite** и **epigraph** этот тэг должен стоять в самом конце родительского тэга, то есть непосредственно перед его (родительского тэга) закрытием. Внутри же роет тэг **<text-author>** може быть либо последний, либо предпоследний перед **<date>**, как это описано в главе ?Стихи? данного описания. *Допускается любое число вхождений тэга <text-author>*.

Вот пример:

```
<epigraph>
  <p>Кажется, знаешь о себе все, так нет.
  Находятся люди, которые знают о тебе больше.</p>
  <text-author>В. Андреев</text-author>
</epigraph>
```

Картинки

Картинки вставляются с помощью тэга **<image/>** (закрывающий тэг не требуется). Он может иметь атрибуты **type**, **href**, **alt**, **title**, **id**. Свойство **href** указывает на адрес картинки. Если адрес картинки начинается с символа # (решетка), то картинка находится внутри документа.

Пример:

```
<p>Абзац текста до картинки.</p>
<image 1:href="#picture.jpg"/>
<p>Абзац текста после картинки.</p>
```

В этом примере в текст вставляется картинка `picture.jpg`, приложенная к документу (на что указывает символ решетки перед именем).

Тэг `` может быть вложен в тэги `<body>`, `<section>`, и в этом случае он описывает просто картинку. Если же `` вложен в `<p>` или `<coverpage>`, он обозначает inline-картинку, то есть такую, которая отображается наравне с текстом, внутри его. У Inline-картинок допустимы только атрибуты **type**, **href**, **alt**.

Файл обложки принято называть `cover.jpg` или `cover.png`.

Ссылки

Ссылки вставляются с помощью тэга `<a>`, который может быть вложен в абзацный тэг `<p>`. Свойство **href** определяет адрес ссылки. Адреса с символа `#` (решетка) задают расположение внутри документа.

Пример:

```
<p>Об этом вы можете прочитать <a href="#n1">здесь</a>.</p>
```

В этом примере текстом ссылки будет `?здесь?`, а указывать она будет на элемент с идентификатором **n1** внутри документа.

В **href** вместо **l** может быть и другой идентификатор, подробнее это описано в главе "Раздел FictionBook".

Также тэг `<a>` может иметь атрибут **type**, что используется при создании ссылок на примечания и описано в разделе "Примечания" данного описания.

Таблицы

Таблицы оформляются при помощи тэга `<table>`, который может быть вложен в `<cite>`, `<section>` и/или `<annotation>`.

Этот тэг может иметь свойства **id** и **style**. **id** описывает идентификатор таблицы, который можно использовать в ссылке, а **style** определяет, какой стиль отображения сопоставлен данной таблице.

Тэг `<table>` содержит вложенные тэги `<tr>...</tr>`, каждый `<tr>` описывает один горизонтальный ряд ячеек таблицы. Этот тэг может иметь атрибут **align**, которое предписывает программе-просмотрщику использовать определенный тип горизонтального выравнивания для ряда таблицы. Этот атрибут может принимать значения **"left"** (по левому краю), **"right"** (по правому краю) или **"center"** (центрирование). По умолчанию используется значение **"left"**.

Тэг `<tr>`, в свою очередь, содержит вложенные тэги `<th>` и `<td>` (в любом количестве и порядке каждый). `<th>` описывает ячейку, которая должна быть отображена как заголовок таблицы. `<tr>` обрамляет простую ячейку таблицы. Эти два тэга могут содержать только текст без всяких тэгов.

Тэги `<th>` и `<td>` могут иметь атрибуты **id**, **style**, **colspan**, **rowspan** и **align**. Свойство **id** задает идентификатор, по которому на ячейку можно сослаться. Атрибут **style** назначает стиль оформления для ячейки (нужно при использовании CSS). Свойства **colspan** и **rowspan** определяют количество колонок и рядов, которые занимает ячейка, что позволяет "растягивать" ее по вертикали или горизонтали. Атрибут **align** описывает тип горизонтального выравнивания, аналогично одноименному свойству тэга `<table>`.

В примере показано, как оформить таблицей небольшую табличку умножения:

```

<table>
  <tr>
    <th></th>
    <th>2</th>
    <th>3</th>
    <th>4</th>
  </tr>
  <tr>
    <th>5</th>
    <td>10</td>
    <td>15</td>
    <td>20</td>
  </tr>
  <tr>
    <th>6</th>
    <td>12</td>
    <td>18</td>
    <td>24</td>
  </tr>
  <tr>
    <th>7</th>
    <td>14</td>
    <td>21</td>
    <td>28</td>
  </tr>
</table>

```

	2	3	4
5	10	15	20
6	12	18	24
7	14	21	28

Примечания

Примечания к тексту книги должны находиться в специальном разделе **<body>**, поименованном **notes**, то есть описан он должен быть как **<body name="notes">**, причем именно так - notes - пятью маленькими буквами и не иначе. Каждое примечание должно находиться в отдельном разделе **<section>** и иметь уникальный **id**, по которому оно будет доступно в ссылках. Вот типичный пример оформления тела примечаний:

```

<body name="notes">
  <title><p>Примечания</p></title>
  <section id="note1">
    <title>
      <p>1</p>
    </title>
    <p>Известный английский архитектор XVIII века, испытавший в
своем творчестве сильное влияние античной архитектуры. (Здесь и
далее прим. ред.)</p>
  </section>
  <section id="note2">
    <title>
      <p>2</p>
    </title>
    <p>Американский просветитель, живший в XVIII веке</p>
  </section>
</body>

```

Ссылка на примечание должна иметь свойство **type**, равное значению **"note"**. Соответствующая размещенному выше примеру ссылка на первое примечание будет выглядеть таким образом:

```

в стиле Роберта Адама <a href="#note1" type="note">[1]</a>

```

Двоичные вложения

К документу FB2 могут быть приложены файлы, например картинки или что-нибудь другое. (Библиотека, правда, не примет файлы с иными вложениями, чем картинки).

Для описания двоичных вложений служит элемент **<binary>**, который идет в документе вслед за последним элементом **<body>**. Вот пример вложения:

```
<binary id="cover.jpg" content-type="image/jpeg">
AQMAEAMDBgkAAA/tAAAE7gAANH/2wCEAAgGBgYGBggGBggMCAcIDA4KCAgKDhANDQ4NDRAR
DA4NDQ4MEQ8SExQTEg8YGBoaGBgjIiIiIycnJycnJycnJycBCQoJCwkJCw4LDQs0EQ40
...
h9x80J/B/c+gfbXM/wDQ/qcPZ9NHjqapi17zZ4+6deEYNxtP+n1v80psec//2Q==
</binary>
```

Свойство **id** описывает идентификатор (имя) двоичного объекта, а свойство **content-type** - MIME-тип вложения. Содержанием элемента **<binary>** является файл, закодированный по методу BASE64.

Чтобы библиотека приняла файл, нужно чтобы в документе не было приложенных объектов, на которые нет ссылки, и чтобы не было ссылок, которые указывают на двоичные объекты, не приложенные к файлу.

Из картинок поддерживаются форматы JPG (тип image/jpeg) и PNG (тип image/png). Поддержка формата GIF отсутствует. Файл обложки обычно принято называть cover.jpg.

Символы Unicode

В текст можно вставлять символы по их коду в таблице Unicode. Для этого нужно написать значок &, потом значок #, потом десятичный код символа, потом точку с запятой.

Например для значка ? (не равно) ? десятичный код 8800, надо написать так:

```
&#8800;
```

Список возможных жанров

Здесь перечислены значения, которые принимает поле **<genre>** в заголовке FB2-документа.

sf_history	Альтернативная история
sf_action	Боевая Фантастика
sf_epic	Эпическая Фантастика
sf_heroic	Героическая фантастика
sf_detective	Детективная Фантастика
sf_cyberpunk	Киберпанк
sf_space	Космическая Фантастика
sf_social	Социальная фантастика
sf_horror	Ужасы и Мистика
sf_humor	Юмористическая фантастика
sf_fantasy	Фэнтези
sf	Научная Фантастика
child_sf	Детская Фантастика
det_classic	Классический Детектив
det_police	Полицейский Детектив
det_action	Боевики
det_irony	Иронический Детектив
det_history	Исторический Детектив
det_espionage	Шпионский Детектив

det_crime	Криминальный Детектив
det_political	Политический Детектив
det_maniac	Маньяки
det_hard	Крутой Детектив
thriller	Триллеры
detective	Детектив
sf_detective	Детективная Фантастика
child_det	Детские Остросюжетные
love_detective	Остросюжетные Любовные Романы
prose	Проза
prose_classic	Классическая Проза
prose_history	Историческая Проза
prose_contemporary	Современная Проза
prose_counter	Контркультура
prose_rus_classic	Русская Классика
prose_su_classics	Советская Классика
humor_prose	Юмористическая Проза
child_prose	Детская Проза
love	Любовные романы
love_contemporary	Современные Любовные Романы
love_history	Исторические Любовные Романы
love_detective	Остросюжетные Любовные Романы
love_short	Короткие Любовные Романы
love_erotica	Эротика
adv_western	Вестерны
adv_history	Исторические Приключения
adv_indian	Приключения: Индейцы
adv_maritime	Морские Приключения
adv_geo	Путешествия и География
adv_animal	Природа и Животные
adventure	Приключения: Прочее
child_adv	Детские Приключения
children	Детское
child_tale	Сказки
child_verse	Детские Стихи
child_prose	Детская Проза
child_sf	Детская Фантастика
child_det	Детские Остросюжетные
child_adv	Детские Приключения
child_education	Детская Образовательная литература
children	Детское: Прочее
poetry	Поэзия
dramaturgy	Драматургия
humor_verse	Юмористические Стихи
child_verse	Детские Стихи
antique_ant	Античная Литература
antique_european	Европейская Старинная Литература
antique_russian	Древнерусская Литература
antique_east	Древневосточная Литература
antique_myths	Мифы. Легенды. Эпос
antique	Старинная Литература: Прочее
sci_history	История
sci_psychology	Психология
sci_culture	Культурология
sci_religion	Религиоведение
sci_philosophy	Философия
sci_politics	Политика
sci_business	Деловая литература
sci_juris	Юриспруденция
sci_linguistic	Языкознание
sci_medicine	Медицина
sci_phys	Физика
sci_math	Математика
sci_chem	Химия
sci_biology	Биология
sci_tech	Технические
science	Научно-образовательная: Прочее
adv_animal	Природа и Животные
comp_www	Интернет
comp_programming	Программирование
comp_hard	Компьютерное Железо
comp_soft	Программы
comp_db	Базы Данных
comp_osnet	ОС и Сети
computers	Компьютеры: Прочее
ref_encyc	Энциклопедии
ref_dict	Словари
ref_ref	Справочники
ref_guide	Руководства
reference	Справочная Литература: Прочее

nonf_biography	Биографии и Мемуары
nonf_publicism	Публицистика
nonf_criticism	Критика
nonfiction	Документальное: Прочее
design	Искусство, Дизайн
adv_animal	Природа и Животные
religion	Религия
religion_rel	Религия
religion_esoterics	Эзотерика
religion_self	Самосовершенствование
religion	Религия и духовность: Прочее
sci_religion	Религиоведение
humor_anecdote	Анекдоты
humor_prose	Юмористическая Проза
humor_verse	Юмористические Стихи
humor	Юмор: Прочее
home_cooking	Кулинария
home_pets	Домашние Животные
home_crafts	Хобби, Ремесла
home_entertain	Развлечения
home_health	Здоровье
home_garden	Сад и Огород
home_diy	Сделай Сам
home_sport	Спорт
home_sex	Эротика, Секс
home	Дом и Семья: Прочее

Список возможных языков

Ниже приведены коды языков, которые можно прописывать в тэгах **<lang>** и **<src-lang>**.

Здесь приведены только основные кодировки, полный список содержится в стандарте ISO 639:

abk	ab	Абхазский
aze	az	Азербайджанский
alb/sqi	sq	Албанский
eng	en	Английский
arm/hye	hy	Армянский
BA		Башкирский
bel	be	Белорусский
bul	bg	Болгарский
hun	hu	Венгерский
vie	vi	Вьетнамский
dut/nla	nl	Голландский
ell/gre	el	Греческий современный (1453-)
dan	da	Данийский
grc		Древнегреческий (до 1453)
heb	he	Иврит
esl/spa	es	Испанский
ita	it	Итальянский
kaz	kk	Казахский
kir	ky	Киргизский
chi/zho	zh	Китайский
kor	ko	Корейский
lat	la	Латинский
lav	lv	Латвийский
lit	lt	Литовский
mac/mak	mk	Македонийский
mol	mo	Молдавский
mon	mn	Монгольский
deu/ger	de	Немецкий
mul		Несколько языков
und		Неопределенный
nor	no	Норвежский
fas/per	fa	Персидский
pol	pl	Польский
por	pt	Португальский
rus	ru	Русский
san	sa	Санскрит
slk/slo	sk	Словакский
slv	sl	Словенский
tgk	tg	Таджикский

tat	tt	Татарский
tur	tr	Турецкий
uzb	uz	Узбекский
ukr	uk	Украинский
cym/wel	cy	Уэльский
fin	fi	Финский
fra/fre	fr	Французский
che		Чеченский
ces/cze	cs	Чешский
	hr	Хорватский
sve/swe	sv	Шведский
epo	eo	Эсперанто
est	et	Эстонский
jpn	ja	Японский

Источник — «http://www.fictionbook.org/index.php?title=Описание_формата_FB2_от_Sclex&oldid=2972»

Категория: FB документы

-
- Последнее изменение этой страницы: 17:38, 27 января 2008.
 - Содержимое доступно по лицензии GNU Free Documentation License (если не указано иное).