

---

# Трансформация FictionBook2 — DocBook — XSL-FO — PDF

KiR

## Содержание

Преобразование FictionBook2 — DocBook .....	1
Описание XSL-трансформации .....	1
Сопоставление FictionBook2 и DocBook .....	3
Скрипт преобразования FictionBook2 — DocBook — XSL-FO — PDF .....	3
Описание .....	3
Использование .....	3
Шаблоны .....	4
Необходимые программные пакеты и библиотеки .....	4
Замечания по генерации PDF .....	5
Download .....	5
Бинарный пакет для Win32 .....	5
Исходный код .....	6
Ссылки .....	6

## Преобразование FictionBook2 — DocBook

### Описание XSL-трансформации

Преобразование FictionBook2 — DocBook написано на XSLT v1.0 с расширениями: элементом `xsl:document` и функциями `document()`, `key()`, `generate-id()`. Данный XSLT-сценарий выполняет следующие действия:

- Переводит FictionBook 2 в DocBook
- Для полей без соответствия в DocBook из `fb:description` пишет соответствующие приложения в конце документа (опционально, см. Таблица 1, «Шаблонные параметры трансформации FictionBook2 — DocBook»);
- Создает индексный файл, в котором находятся:
  - список файлов с их `mime`-типом;
  - метаинформация о документе (авторы, название);
  - размеры и разрешение (в dpi) вывода.
- Выводит все вложения в `.fb2` (тег `binary`) в отдельные файлы

**Об интернационализации.** В `fb2docbook/l10n` находятся файлы для поддержки языков (сейчас русский и английский). Если файл под какой-то язык отсутствует, то выбирается язык по умолчанию (см. Таблица 1, «Шаблонные параметры трансформации FictionBook2 — DocBook»).

**Входные параметры для xslt-трансформации.**

1. `conv_info_idx` — индексный файл. См. transformation output index [1].
2. `document-element` — название элемента типа `xsl:document`. Зависит от используемого процессора.

**Шаблонные параметры трансформации.** Находятся в шаблонах параметров, используемых для конвертации DocBook — XSL-FO, см. Таблица 5, «Шаблоны основных страниц». Одинаковы для всех шаблонов (кроме начинающихся с *output.*), см. включаемый файл `params/p_fb2_common-default.xml`.



### Внимание

Обязательно прочитайте про параметры *output.\**, *page.width* и *page.height*. В скрипте `btransformer.py` пока обрабатываются только такие варианты записи.

**Таблица 1. Шаблонные параметры трансформации FictionBook2 — DocBook**

Название параметра	Описание
<i>fb2.book-parts</i>	Указывает как вложенные <code>fb:section</code> должны заменяться на соответствующие структурные элементы DocBook (например, на первом (верхнем) уровне -- это <code>part</code> , а идущие на том же уровне <code>fb:epigraph</code> и <code>fb:image</code> идут в <code>partintro</code> ).
<i>fb2.image.align</i>	Горизонтальное расположение изображения (в FB2 такое указание для рисунков отсутствует).
<i>fb2.image.valign</i>	Вертикальное расположение изображения (в FB2 такое указание для рисунков отсутствует).
<i>fb2.subtitle.role</i>	Выделение (жирный, курсив и т.д.) для тега <code>fb:subtitle</code> .
<i>fb2.default.language</i>	Язык по умолчанию, если он не указан в FB2-документе или отсутствует его интернационализационная поддержка.
<i>fb2.print.infos</i>	Булев параметр, выводить или нет информацию в виде приложений о FB2-документе.

**Таблица 2. Шаблонные параметры трансформации FictionBook2 — DocBook, только записываемые в индексный файл**

Название параметра	Описание
<i>output.images_mode.resize</i>	Булев параметр. Определяет, преобразовывать ли изображения или нет. Возможные варианты записи: 1, 0, True, true, False, false.  Если истинен, обязательно определять все нижеследующие параметры!
<i>output.dpi.width</i>	Количество точек на дюйм по горизонтали.  Записывается просто числом в строке!
<i>output.dpi.height</i>	Количество точек на дюйм по вертикали.  Записывается просто числом в строке!
<i>output.max_image_margin.width</i>	Если умножить на 2 и полученное число вычесть из ширины страницы в пикселях, то получим максимальный горизонтальный размер изображения.  В пикселях! Записывается просто числом в строке!
<i>output.max_image_margin.height</i>	Если умножить на 2 и полученное число вычесть из высоты страницы в пикселях, то получим максимальный вертикальный размер изображения.  В пикселях! Записывается просто числом в строке!

**Таблица 3. Шаблонные параметры трансформации FictionBook2 — DocBook, только записываемые в индексный файл, и общие с параметрами для DocBook**

Название параметра	Описание
<code>page.width</code>	Ширина страницы.  В миллиметрах. Записывается как действительное число (разделитель — точка), далее без пробела 'mm'. Например, '90.6mm'
<code>page.height</code>	Высота страницы.  В миллиметрах. Записывается как действительное число (разделитель — точка), далее без пробела 'mm'. Например, '122.4mm'

## Сопоставление FictionBook2 и DocBook

TODO

# Скрипт преобразования FictionBook2 — DocBook — XSL-FO — PDF

## Описание

Скрипт написан на Python. Служит "запускалкой" для xsl-трансформаций и других инструментов, скрывая детали низкоуровневой реализации. Переводит вложенные в .fb2 файлы обратно в бинарный вид и выполняет трансформацию изображений (изменение размера и цветности).

## Использование



### Замечание

Входной .fb2-файл может быть сжат:

- **gzip** (расширение .gz)
- **bzip2** (расширение .bz2)
- **zip** (расширение .zip), в данном случае обработается файл в архиве с именем, совпадающим с названием архива, но без суффикса '.zip'

```
usage: btransformer.py --config config_file [options] fb2_file pdf_file
```

options:

```
--version          show program's version number and exit
-h, --help         show this help message and exit
-c FILE, --config=FILE read config from FILE
--wh              turn on win32hack -- needed for using on win32
                  without tools installation
--no-remove_tmp    do not remove working temporal files
--xslt_proc=XSLT_PROC set xslt_proc to use
--fo_proc=FO_PROC   set fo_proc to use
--pdf_proc=PDF_PROC set pdf_proc to use
--xslt_params=FILE  set xslt_params file to use
--xslt_titlepages=FILE set xslt_titlepages file to use
```

## Шаблоны

В пакеты входят шаблоны, приспособленные для вывода pdf из формата FB2 и для устройства Sony Reader. Находятся в директории `params` дистрибутива.

**Таблица 4. Шаблоны для заглавных страниц**

Название файла шаблона	Формат или устройство	Примечание
<code>titlepage.templates_prs-default.xml</code>	Sony Reader PRS-500	Уменьшены размеры шрифтов заглавий
<code>titlepage.templates_a4-default.xml</code>	Бумага, размер A4	Изменений не вносилось (копия из дистрибутива Docbook XSL Stylesheets)
<code>titlepage.templates_a5-default.xml</code>	Бумага, размер A5	Для распечатки по две страницы на A4  Изменений не вносилось (копия из дистрибутива Docbook XSL Stylesheets)

**Таблица 5. Шаблоны основных страниц<sup>a</sup>**

Название файла шаблона	Формат или устройство	Шрифты	Примечание
<code>params_prs-font.liberation.xml</code>	Sony Reader PRS-500	Шрифты Liberation [ <a href="https://www.redhat.com/promo/fonts/">https://www.redhat.com/promo/fonts/</a> ]	
<code>params_prs-font.vista.xml</code>	Sony Reader PRS-500	Шрифты из Vista  основной — Cambria без засечек — Candara моноширинный — Consolas	
<code>params_a4-font.liberation.xml</code>	Бумага, размер A4	Шрифты Liberation [ <a href="https://www.redhat.com/promo/fonts/">https://www.redhat.com/promo/fonts/</a> ]	
<code>params_a4-font.vista.xml</code>	Бумага, размер A4	Шрифты из Vista  основной — Cambria без засечек — Candara моноширинный — Consolas	
<code>params_a5-font.liberation.xml</code>	Бумага, размер A5	Шрифты Liberation [ <a href="https://www.redhat.com/promo/fonts/">https://www.redhat.com/promo/fonts/</a> ]	Для распечатки по две страницы на A4
<code>params_a5-font.vista.xml</code>	Бумага, размер A5	Шрифты из Vista  основной — Cambria без засечек — Candara моноширинный — Consolas	Для распечатки по две страницы на A4

<sup>a</sup>Во всех шаблонах параметры для преобразования FictionBook2 — DocBook одинаковы.

## Необходимые программные пакеты и библиотеки

- XSLT-процессор, поддерживающий спецификацию XSLT v1.0 и инструкцию типа `xsl:document` из спецификации XSLT v1.1. Один из:

- Libxslt [<http://xmlsoft.org/XSLT/>] (как в виде модуля под питон, так и в виде отдельной программы **xsltproc**)
- Saxon v6 [<http://saxon.sourceforge.net/>]
- 4suite [<http://4suite.org/>]
- Xalan [<http://xml.apache.org/xalan-j/index.html>] <sup>1</sup>
- XSL Stylesheets for Docbook [<http://wiki.docbook.org/topic/DocBookXslStylesheets>] v1.72.0 и выше
- XSL-FO-процессор. Один из:<sup>2</sup>
  - FOP [<http://xmlgraphics.apache.org/fop/>] v0.93
  - XEP [<http://www.renderx.com/>]
- Python [<http://www.python.org>] v2.4 и выше
- Python imaging library (PIL) [<http://www.pythonware.com/products/pil/index.htm>]
- Pdftk [<http://www.pdfhacks.com/pdftk>]

## Замечания по генерации PDF

**Генерация в FOP.** FOP пишет в .pdf только содержание (bookmarks), другую метаинформацию (автор, название) приходится дописывать с помощью pdftk. К сожалению в таком случае Acrobat Pro/Acrobat Reader (по-крайне мере версии 8) показывает только поле Title (название), при этом вьюер в Sony Reader показывает все поля.

**Генерация в XEP.** XEP имеет расширение к XSL-FO, которое позволяет дописывать всю метаинформацию, и в .pdf, полученных с использованием этого процессора, в Adobe Acrobat виден весь description (описание).

**Линеаризация.** Для комфортного просмотра pdf-документов при медленном доступе к файлу или при малом количестве оперативной памяти, необходимо проводить линеаризацию документа (так называемый 'fast web view' в Acrobat). Для этого можно использовать либо Adobe Acrobat (File - Save as (Adobe pdf files, optimized, \*.pdf) и выставить соответствующую опцию в Settings), либо скрипт **pdfopt** из пакета Ghostscript [<http://en.wikipedia.org/wiki/Ghostscript>]. Но если использовать **pdfopt**, то просмотрщик в PRS-500 не показывает страницы после оглавления вообще, они просто пусты (сложно сказать где и у кого в данном случае ошибка или недоработка). В случае же оптимизации в Acrobat'e, полученные документы быстро отображаются и листаются на Sony Reader на любой и с любой страницы.

## Download

### Бинарный пакет для Win32

**Download.**

- fb2d.pkg.bin.win32-0.1.bzr236.zip  
[<http://kir666.ru/fb2docbook/download/fb2d.pkg.bin.win32-0.1.bzr236.zip>]

**Требования.** Установленная JVM.

**Установка и запуск.** Распаковать в директорию по выбору. Для запуска использовать скрипт **btransformer.cmd**.

---

<sup>1</sup>Не поддерживает вывод вложенных файлов, т.к. xalan:write всегда выводит в выходном файле xml-заголовок.

<sup>2</sup>Оба процессора написаны на Java и требуют установленной JVM



### Состав пакета.

- ## Исходный код

• f b 2 d o c b o o k . s r c - 0 . 1 . b z r 2 3 6 . z i p  
[<http://kir666.ru/fb2docbook/download/fb2docbook.src-0.1.bzr236.zip>]

## Ссылки

- <sup>3</sup>Метрики для FOP'a в архиве присутствуют.