

Секреты "Apache Forrest"

Table of contents

1 Введение.....	2
2 Настройка русских шрифтов для генерации pdf.....	2
3 Фичи.....	3

1. Введение

Сейчас создавать сайты на "голом" HTML'e бессмысленно, т.к. уже давно изобретены разные инструменты для их быстрого создания, где вас будет только заботить информационная часть, изредка программная. Но, большинству движков необходим какой-либо скриптовый язык программирования, а так же, базы данных. Но большинство бесплатных хостингов ограничиваются только показом HTML, поэтому я начал искать инструмент для создания статических сайтов. И он был найден. Имя ему - "**Apache Forrest**". Стоит упомянуть, что данный сайт сделан именно на этом движке. Более подробную информацию вы можете получить на [сайте проекта](http://forrest.apache.org/) (<http://forrest.apache.org/>) и на - [Javable](http://www.javable.com/about/forrest) (<http://www.javable.com/about/forrest>) . Скажу лишь, что для создания сайта вам потребуется JVM, знание XML и, если вы будете менять темы - минимальные знания xsl, css, html. Для редактирования xml-файлов, я пользуюсь [emacs'ом](http://www.gnu.org/software/emacs/) ([../linux/emacs.html](http://www.gnu.org/software/emacs/)) , чего и вам советую.

2. Настройка русских шрифтов для генерации pdf

Из коробки, forrest, по крайней мере версия 0.6, в pdfках вместо русского текста, генерит решетки "#". Казус весь в настройках шрифтов для [fop](http://xmlgraphics.apache.org/fop/) (<http://xmlgraphics.apache.org/fop/>) . Озадачившись этой проблемой, я ринулся в гугл, который меня вывел на сайт [Javable](http://www.javable.com/about/forrest) (<http://www.javable.com/about/forrest>) , где эта проблема была решена в дистрибутиве метного разлива. Скажем спасибо этим добрым людям или доброму человеку.

Далее будет упомянуться пути на каталоги, где:

\$FORREST_HOME - путь до директории, где установлен форрест

\$FORREST_CONTEXT - путь, где лежат глобальные конфигурационные файлы, к примеру, "\$FORREST_HOME/src/core/context"

Нам необходимо будет [сгенерировать метрики для шрифтов](http://xmlgraphics.apache.org/fop/fonts.html#truetype-metrics)

(<http://xmlgraphics.apache.org/fop/fonts.html#truetype-metrics>) или взять уже [готовые](http://www.gnu.org/software/emacs/) ([../files/forrest/fontmetrics.tar.bz2](http://www.gnu.org/software/emacs/)) , положить их, к примеру, в каталог "\$FORREST_CONTEXT/fontmetrics". Сами шрифты можно будет установить в "\$FORREST_CONTEXT/fonts".

Прописываем в "\$FORREST_CONTEXT/sitemap.xmap", что используем конфиг для fop:

Секреты "Apache Forrest"

```
...
  <map:serializer name="fo2pdf"
    src="org.apache.cocoon.serialization.FOPSerializer"
    mime-type="application/pdf">
    <user-config src="userconfig.xml"/>
  </map:serializer>
...
```

Создаем (копируем) конфиг [userconfig.xml](#) (./files/forrest/userconfig.xml.tar.bz2) в папку "\$FORREST_CONTEXT". В этом файле сопоставляем шрифт и метрику. Я его описывать не буду, т.к. описание есть на сайте [fop](#) (<http://xmlgraphics.apache.org/fop/configuration.html>). Шрифты, описанные в этом самом конфиге (в моем примере), копируем в "\$FORREST_CONTEXT/fonts".

Затем, чтобы содержание в начале отображалось на русском языке, меняем файл "\$FORREST_CONTEXT/fonts/skins/common/xslt/fo/document2fo.xml":

```
..
<xsl:template match="body[count(//section) != 0]">
<xsl:if test="$toc-max-depth > 0">
  <fo:block font-family="serif" font-size="14pt" font-weight="bold"
    space-after="5pt" space-before="5pt" text-align="justify"
width="7.5in">
  <!-- insert i18n stuff here -->
  <xsl:text>Table of contents</xsl:text>
</fo:block>
-- <fo:block font-family="sans" font-size="12pt" space-after="5pt"
++ <fo:block font-family="serif" font-size="12pt" space-after="5pt"
  space-before="0pt" text-align="justify" width="7.5in">
  <xsl:if test="$page-break-top-sections">
    <xsl:attribute name="break-after">page</xsl:attribute>
  </xsl:if>
  <xsl:apply-templates select="section" mode="toc" />
</fo:block>
</xsl:if>
<xsl:apply-templates />
</xsl:template>
..
```

Вот и все. Теперь в pdfках можно лицезреть русские буквы.

3. Фичи

На хостинге www.narod.ru (<http://www.narod.ru>) HTML должен "отдаваться" в кодировке windows-1251 (forrest по-умолчанию создает хтмл в юникоде). Для этого меняем в файле "\$FORREST_CONTEXT/sitemap.xmap":

```
...
  <map:serializer name="html" mime-type="text/html"
src="org.apache.cocoon.serialization.HTMLSerializer">
  <doctype-public>-//W3C//DTD HTML 4.01
Transitional//EN</doctype-public>
<doctype-system>http://www.w3.org/TR/html4/loose.dtd</doctype-system>
--   <encoding>UTF-8</encoding>
++   <encoding>WINDOWS-1251</encoding>
  </map:serializer>
...
```

А фича вот в чем. В самом html-коде, вы будете видеть все в кодах символах, к примеру, &0012; и т.д. :). Я сильно парится не стал, слабал скрипт, который все хтмл-файлы перед отправкой на narod.ru перекодирует из юникода в windows-1251. Хотя, можно и оставить, т.к. сам файл будет соответствовать стандартам.