

Perl だけで作る日本語 PDF

日本語組版ルールを組み込んだ PDF 生成モジュール PDFJ

中島 靖 <nakajima@netstock.co.jp>

2002/5/11

Kansan.pm 2 周年イベント

背景と目的 ~ 1

Web アプリケーションに印刷機能を持たせるために
サーバーで動的に日本語 PDF を生成したい

TeX+dvipdfm	<ul style="list-style-type: none">* サーバー側にたくさんのソフトをインストールしないといけない* 複雑な数式を扱うならこれしかないが、ビジネスアプリには不要だろう
PDFLib	<ul style="list-style-type: none">* 日本語対応の細かなところまでは無理* 商用の場合ライセンスが必要
FOP	<ul style="list-style-type: none">* うーん、JAVA かあ...(^;::)* 日本語対応に難があるらしい... (よく知らない、ごめんなさい)

Perl だけでちょいちょいっとやりたいなあ

背景と目的 ~ 2

本を書くのに LaTeX のマクロの代わりに Perl スクリプトでレイアウトできたらラクだなあ

LaTeX で複雑なマクロを書いたら不安定で困った...(すみません、TeXnician じゃないもので)

自作モジュールなら日本語組版ルールもきちんと組み込めるだろう

TeX では難しい、ルビの配置とか、追い込みの時の空白の詰め方とか...(すみません、TeXnician じゃないもので)

目標

- * Perl だけで書く（C のコンパイルができないサーバーでも使えるように）
- * 日本語組版ルール（JIS X 4051）を組み込む
- * 縦書きも OK
- * 段落、表、箇条書き、図といったレイアウト要素のページへの適切な配置
- * アウトラインやハイパーリンクといった PDF 機能も利用できるようにする
- * できる限りフリーにする

日本語組版ルール～ 1

JIS X 4051「日本語文書の行組版方法」(1995)を採用

JIS X 4051のうち次のものをサポート

- * 約物の幅と間隔
- * 禁則処理と分離禁止処理
- * 縦中横
- * ルビ
- * 添え字
- * 圏点
- * 下線・傍線
- * 行長揃えのための間隔調整方法
- * 欧文のハイフネーション

日本語組版ルール～ 2

JIS X 4051 と異なる点

- * 文字間の間隔はすべて後ろの文字のサイズに従う
- * 欧文スペースは三分空きでなく半角空き
- * 割注は未サポート
- * タブ処理は未サポート

入力はどうする？

Perl スクリプトで直接書くのでない限り、なんらかのマークアップ言語を入力とする必要

独自のマークアップ言語	<ul style="list-style-type: none">* 必要なものだけ実装できる* 自分で本を書くためならこれがラク
LaTeX のサブセット	<ul style="list-style-type: none">* LaTeX からの移行がラク
XSL	<ul style="list-style-type: none">* 将来性として有望* フルサポートは大変、サブセットか
HTML+CSS	<ul style="list-style-type: none">* 既存文書をそのまま使える* フルサポートは大変、互換性の泥沼が...
その他	<ul style="list-style-type: none">* pod...pod2pdf はそれなりに便利かも

現状と計画

未完成です

未実装の主なもの

- * レイアウト要素のページへの自動分割・配置
- * アウトラインとハイパーリンク
- * マークアップ言語による入力
- * 説明書

説明書を書いて バージョンとして公開

デモ

「約物の、(幅と間隔)」

ゴシック Gothic 明朝 Minchou
ゴシック Hankaku 明朝 Hankaku
色文字 下線 下破線 圏点 網掛け 日本語斜体
添え字¹⁾ 中島 文字位置
あけぼの 曙 大親分^{ボ ス} 大親分^{boss}
フィフティーン fifteen

大親分^{boss} fifteen
フィフティーン

添え字⁽¹⁾ 中島^{なかじま} 文字位置^{もじいち} 曙^{あけぼの} 大親分^{ボ ス}
圏点、網掛け 日本語斜体 (1) 縦中横
Hankaku 色文字 傍線 傍破線

ゴシック Gothic 明朝 Minchou
シツク Hankaku 明朝