

Zpravodaj Československého sdružení uživatelů TeXu

Pavel Stríž

Představení formátu XeLaTeX

Zpravodaj Československého sdružení uživatelů TeXu, Vol. 20 (2010), No. 4, 292–296

Persistent URL: <http://dml.cz/dmlcz/150146>

Terms of use:

© Československé sdružení uživatelů TeXu, 2010

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ*:
The Czech Digital Mathematics Library <http://dml.cz>

Abstrakt

Novinkou posledních několika let ve světě T_EXu jsou rozšíření nazvané X_YT_EX (<http://www.tug.org/xetex/>) a LuaT_EX (<http://www.luatex.org/>). V následujících odstavcích si představíme formát X_YLaT_EX, neboť jeho dokumentace je úplnější a pro začátečníka to je vstřebatelnější než rozšíření LuaT_EX. Autorem X_YT_EXu je Jonathan Kew. Tvůrce nosného podpůrného T_EXového balíčku nazvaného fontspec je Will Robertson.

Klíčová slova: X_YT_EX, X_YLaT_EX, UTF-8, japonština, písmo, Code2000, TTF, OTF, balíček fontspec, balíček otinfo.

doi: 10.5300/2010-4/292

Předpoklady

Slabinou T_EXu bývalo, že nebylo bez konverzí do PFB+TFM možné přímo používat písma ve formátu TTF a novějším formátu OTF. Jakmile narazíme na sazbu ČJKV (čínština, japonština, korejština, vietnamština), tak i s konverzemi u nově staženého písma můžeme předčasně neúspěšné práce ukončit. I když třeba známe práci s balíčkem CJK, či rozšířeními C_T_EX [3], ko_T_EX či p_T_EX [1], tak to stále neřeší náš problém, jak použít hezké písmo v TTF/OTF, kde jsou znaky třeba jen v rozsahu latinky. Tyto slabiny zdárně překonávají X_YT_EX [2] i LuaT_EX [4], které navíc umí přímou práci s kódováním dokumentů v UTF-8.

Hello World! alias 今日は！

Za předpokladu, že máme k dispozici některé systémové písmo, např. Verdanu, a nainstalován T_EX Live 2010, tak jsme plně vybaveni na první experimenty.

Naše základní ukázka by mohla být zakotvena v souboru `prvni.tex`:

```
% xelatex prvni.tex
\documentclass{article}
\usepackage{fontspec}
\begin{document}
Nazdar světe!\par
\fontspec{Verdana}
Nazdar světe!
\end{document}
```

Příprava článku byla podpořena projektem ESF reg. č. CZ.1.07/2.2.00/07.0361.

V příkazové řádce zadáváme `xelatex prvni.tex` a potvrdíme klávesou Enter. V editorech stylu \TeX works si navolíme „XeLaTeX“. V této minimální ukázce se zdrojový kód přepne ze základního písma na systémové písmo Verdana a vypíšeme tuto větu do souboru `prvni.pdf`.

Pokud nemáme písmo Verdana¹, můžeme zkusit DejaVu², Linux Libertine³, jednoduše takové, které máme aktuálně k dispozici.

Nazdar světe!
Nazdar světe!

Kvůli obecnosti si stáhneme písmo Sazanami Mincho⁴ [přímý odkaz⁵] a také písmo Code2000⁶ [přímý odkaz⁷]. Soubor TTF si rozbalíme a uložíme do adresáře, kde vznikají naše \TeX ové soubory. Naše druhá, pokročilejší ukázka v podobě souboru `hello.tex` by mohla vypadat následujícím způsobem:

```
% xelatex hello.tex
\documentclass{article}
\usepackage{fontspec}
\begin{document}
Hello World!\par
\fontsize{20}{24}
\fontspec[ExternalLocation=.,Script=Latin]{CODE2000.TTF}
Hello World!\par
\fontspec[ExternalLocation=.,Script=Kana]{sazanami-mincho.ttf}
今日は！\par
\normalsize\normalfont
Další partie textu\ldots
\end{document}
```

Vedle minimální struktury \LaTeX ového dokumentu si načteme balíček `fontspec`. První a poslední řádek textu je běžné písmo. Příkaz `\fontsize` nám nastaví velikost písma (20pt) a výšku dvou po sobě jdoucích účaří (24pt). Klíčový je příkaz `\fontspec`, který přepíná písmo. Pozdrav v japonštině „Dobrý den!“ (zápis z klávesnice je Kon'nichiha!, čten je koničiva) byl zapsán z klávesnice přímo a následně byl převeden na kandži⁸. Nebo si vykopírujte kandžiky ze slovníku z překladače⁹. Výstup by měl dopadnout takto:

¹<http://en.wikipedia.org/wiki/Verdana>

²http://dejavu-fonts.org/wiki/Main_Page

³<http://www.linuxlibertine.org/>

⁴<http://packages.debian.org/sid/ttf-sazanami-mincho>

⁵<http://sourceforge.jp/projects/efont/downloads/10087/sazanami-20040629.tar.bz2/>

⁶http://www.code2000.net/code2000_page.htm

⁷<http://code2000.net/CODE2000.ZIP>

⁸<http://www.sljfaq.org/afaq/input-methods.html>

⁹<http://translate.google.com/#en|ja|good%20day>



Potenciál písem

Pokud bychom si přáli seznámit se s možnostmi jednotlivých písem, tak lze využít nástroj `otfinfo` (součást `TEX Live`). Možnosti nástroje získáme pomocí `otfinfo --help`. Konkrétně si můžeme zkusit: `otfinfo -s CODE2000.TTF`.

Především novodobá písma mohou obsahovat celou škálu vlastností (scripts, features, optical sizes). Platí to především na přepínače kapitálek, alternativní kresby znaků, ligatur, cifer různých stylů, optického škálování atp. Stojí za pozornost nahlédnout na manuál balíčku `fontspec`¹⁰ a `XYTEXu`¹¹. Ty lze nejrychleji vyvolat z příkazové řádky pomocí `texdoc fontspec`, respektive `texdoc xetex`. Nejkompletněji popisuje `XYTEX` kapitola¹² z knihy *The L^AT_EX Graphics Companion* [2] a především *The X_ET_EX Companion*¹³.

Cvičné OTF písmo s mnoha vlastnostmi připravil Adam Twardoch pod názvem Nadyezhda SL One¹⁴ [přímý odkaz¹⁵]. Řadu písem lze nalézt na stránkách SIL International¹⁶, na testovací účely lze doporučit webové stránky 1001 Free Fonts¹⁷ a jejich galerie písem.

Na rychlé náhledy vlastností OTF písem lze použít Fontmatrix¹⁸.

Devět způsobů zápisu jednoho glyfu

Ne každý bude mít zájem si instalovat čínskou či japonskou klávesnici kvůli několika glyfům. Výhoda práce s UTF-8 kódováním je, že si lze glyfy vykopírovat z webových stránek, projektů Wikimedia, elektronických dokumentů, překladače Google¹⁹ (zkuste si konkrétně kandži pro rybu²⁰) či některých elektronických slovníků (Ctrl+C a Ctrl+V). Také je k ruce Mapa znaků.

¹⁰<http://www.ctan.org/tex-archive/macros/xetex/latex/fontspec/fontspec.pdf>

¹¹<http://ftp.cstug.cz/pub/tex/CTAN/info/xetexref/XeTeX-reference.pdf>

¹²<http://xml.web.cern.ch/XML/lgc2/tlgc2extra.pdf>

¹³<http://xml.web.cern.ch/XML/lgc2/xetexmain.pdf>

¹⁴<http://www.twardoch.com/fonts/>

¹⁵<http://www.twardoch.com/fonts/NadyezhdaSLOne.zip>

¹⁶<http://scripts.sil.org/>

¹⁷<http://www.1001freefonts.com/>

¹⁸<http://fontmatrix.net/>

¹⁹<http://translate.google.com/>

²⁰<http://translate.google.com/#en|ja|fish>

Můžeme se však opřít o \XeTeX a vyzkoušet si vstup dalšími alternativními způsoby, pokud o daném glyfu víme nebo zjistíme dodatečné informace. V získávání podrobností o glyfech přímo ze souboru písem může být nápomocen program FontForge²¹. Umí pracovat v grafickém režimu, dávkově i jako knihovna²² jazyka Python.

```
% xelatex devitka.tex
\documentclass{article}
\usepackage{fontspec}
\begin{document}
\fontspec[ExternalLocation]{CODE2000.TTF}
魚% Přímé vložení znaku.
~~~~9b5a% Zápis se stříškami.
\char"9B5A% Zápis v šestnáctkové soustavě.
\char'115532% Zápis v osmičkové soustavě.
\char39770% Zápis v desítkové soustavě.
\symbol{"9B5A}% Alternativní příkaz.
\XeTeXglyph\the\XeTeXglyphindex"uni9B5A"% Název glyfu.
\XeTeXglyph\the\XeTeXcharglyph"9B5A%
\XeTeXglyph36294% Index glyfu.
\end{document}
```

Výstupem získáváme devětkrát kandži znamenající v japonštině ryba. Ta podoba se nezapře, že?

魚魚魚魚魚魚魚魚

Na rozloučenou

Rozloučíme se s vámi vánočním písmem Bodie MF Holly²³ [přímý odkaz²⁴] takto:

```
% xelatex nashle.tex
\documentclass{article}
\usepackage{fontspec}
\begin{document}
\fontsize{20}{24}
\fontspec[ExternalLocation=.] {BodieMFHolly.ttf}
Na shledanou!
\end{document}
```

²¹<http://fontforge.sourceforge.net/>

²²<http://fontforge.sourceforge.net/python.html>

²³<http://www.1001freefonts.com/christmas-fonts.php>

²⁴<http://www.1001freefonts.com/font/BodieMFHolly.zip>



Seznam použité a doporučené literatury

- [1] Eyre, Timothy. Typesetting Japanese with pTeX. [Sazba japonštiny pomocí pTeXu.] In *Zpravodaj Československého sdružení uživatelů T_EXu*, Volume 20, Number 3, pp. 152–173, 2010. ISSN 1211-6661. doi:10.5300/2010-3/152
- [2] Goossens, Michel; Mittelbach, Frank; Rahtz, Sebastian; Roegel, Denis; Voß, Herbert. *The L^AT_EX Graphics Companion*. [Velký průvodce grafikou v L^AT_EXu.] 2. vyd. USA, Boston: Addison-Wesley Professional, 2007. ISBN 978-0-321-50892-8.
- [3] Jiang, Jjgod. Chinese T_EX Typesetting: Past and Present. [Sazba čínštiny v T_EXu: historie a současnost.] In *Zpravodaj Československého sdružení uživatelů T_EXu*, Volume 20, Number 3, pp. 215–219, 2010. ISSN 1211-6661. doi:10.5300/2010-3/215
- [4] Mádr, Michal; Stríž, Pavel. Představení LuaT_EXu. [Introduction to LuaT_EX.] In *Zpravodaj Československého sdružení uživatelů T_EXu*, Volume 19, Number 4, pp. 191–200, 2009. ISSN 1211-6661. doi:10.5300/2009-4/191

Summary: Introduction to the X_ƎL^AT_EX format

The article warmly welcomes us to the world of X_ƎT_EX (<http://www.tug.org/xetex/>), especially to the X_ƎL^AT_EX format which is common format these days next to LuaL^AT_EX. X_ƎT_EX supports processing UTF-8 coded documents and direct TTF and OTF font loading and use. The author presents several very basic examples accompanied by the output previews.

Key words: X_ƎT_EX, X_ƎL^AT_EX, UTF-8, Japanese, Font, TTF, OTF, Code2000, fontspec package, otfinfo package.

*Pavel Stríž, striz@fame.utb.cz
Tomas Bata University in Zlín
nám. T. G. Masaryka 5555
76001 Zlín, Czech Republic*