

# Zpravodaj Československého sdružení uživatelů TeXu

---

Jan Přichystal; Jiří Rybička

Webové rozhraní pro sazbu dokumentů

*Zpravodaj Československého sdružení uživatelů TeXu*, Vol. 14 (2004), No. 3-4, 190–195

Persistent URL: <http://dml.cz/dmlcz/149972>

## Terms of use:

© Československé sdružení uživatelů TeXu, 2004

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>

---

---

# Webové rozhraní pro sazbu dokumentů

JAN PŘICHYSTAL, JIŘÍ RYBIČKA

Jedním z nedostatků systému  $\text{T}_{\text{E}}\text{X}$  je pomalá a komplikovaná instalace a konfigurace. Tento fakt někdy způsobuje nechuť začínajících uživatelů se s tímto systémem seznámit a pracovat s ním. Tento problém je dobře znám a byl diskutován i ve Zpravodaji CSTUG v anketě uživatelů  $\text{T}_{\text{E}}\text{X}$ u (Polách, 2003). Bylo zde zmíněno, že tento problém by pomohlo vyřešit vytvoření webového rozhraní, které by simulovalo funkčnost  $\text{T}_{\text{E}}\text{X}$ u. Popisovaný projekt „ $\text{T}_{\text{E}}\text{X}$ onWeb“ se snaží tyto nedostatky překonat a nabídnout prostředí pro představení vlastností sázecího systému bez nutnosti jej instalovat na vlastní počítač.

Následně lze tohoto webového stroje použít i pro zpřístupnění určitých typů dokumentů (dopisy, objednávky apod.) běžným uživatelům, přičemž centralizovaná správa pomůže zajistit správnost podle příslušných norem a doporučení.

## Úvod

Netriviální instalace a konfigurace systému  $\text{T}_{\text{E}}\text{X}$  způsobuje začínajícím uživatelům nemalé potíže. Ti často nejsou schopni instalaci tohoto systému zvládnout sami a potřebují pomoc odborníka, na tento systém zanevrou a raději se zase vrátí k programům méně kvalitním, avšak z jejich pohledu jednodušším. Přestože je již delší dobu k dispozici poměrně propracovaná distribuce  $\text{T}_{\text{E}}\text{X}$ Live, setkáváme se často s různými problémy, jejichž společnou příčinou je značná různost operačních prostředí, v nichž se instalace provádí. Podle našich zkušeností se začátečníky je zhruba polovina z nich nucena využívat při instalaci distribuce  $\text{T}_{\text{E}}\text{X}$ Live různých forem technické pomoci. Uvedený podíl je značný a nelze do budoucna předpokládat jeho dramatické snížení, bude-li i nadále každá distribuce koncipována jako velmi komplexní, propracovaný a vybavený systém využitelný jen hrstkou skutečných odborníků.

Pod dojmem této skutečnosti bylo navrženo zpracování projektu webového portálu simulujícího funkčnost systému  $\text{T}_{\text{E}}\text{X}$ , který by začínajícím uživatelům nabídl prostředí systému  $\text{T}_{\text{E}}\text{X}$ , aniž by jej bylo nutné instalovat, a který by byl snadno dostupný a použitelný kdykoliv. Přístup přes webové rozhraní byl zvolen z několika důvodů. Jedním je výhoda vysoké dostupnosti z libovolného místa a operačního systému, kde je přístup k Internetu. Další výhodou je možnost přístupu v jakémkoliv čase, neboť webový server bude neustále v provozu.

Bylo nutné se také zamyslet nad tím, co by takovýto systém měl uživateli konkrétně nabízet, neboť je jasné, že stavu, aby byl naprosto srovnatelný s lokální instalací, nelze v podmínkách webového prohlížeče dosáhnout. Principiálně chybí například veškeré vymoženosti textového editoru, jako je zvýrazňování syntaxe nebo funkce vyhledávání a nahrazování řetězců v textu. Chybí rovněž interaktivní řízení překladu. Uživatel tedy bude mít možnost zapisovat zdrojový text, připojovat soubory dle potřeby a v omezené míře nastavovat způsob překladu.

V zásadě jsme se soustředili na dva případy:

- Uživatel chce napsat pouze kratší text (například představit vlastnosti systému  $\text{T}_{\text{E}}\text{X}$  na několika krátkých příkladech). V takové situaci mu nepohodlí pořizování textu v prohlížeči příliš nepřekáží. Bude potřebovat prostředek pro zapisování zdrojového textu, který bude suplovat funkci textového editoru. Ve spojení s tím může vyvstat požadavek na vkládání vlastních souborů.
- Pro uživatele není problém pořídit zdrojový text na vlastním počítači, nebo ho již dokonce má hotový, a chce provést pouze jednorázový překlad  $\text{T}_{\text{E}}\text{X}$ em na počítači, kde tato možnost není. Pravděpodobně je však k dispozici nějaký textový editor, ve kterém text může napsat a připojení k Internetu není v dnešní době nijak výjimečné.

Dalším požadavkem je možnost vytvářet dokumenty jak formátu PostScript tak PDF, což běžná instalace  $\text{T}_{\text{E}}\text{X}$ u umožňuje. Je také potřeba ovlivňovat počet průchodů překladače. V některých případech totiž nestačí průchod jeden. Typickým příkladem požadavku na větší počet průchodů je sazba obsahu, kdy se musí informace o názvech kapitol a číslech stránek přenést do místa, kde se objeví obsah. Tento problém lze řešit minimálně dvojím překladem: při prvním průchodu se potřebné informace zaznamenají do pomocného souboru, při druhém průchodu se tyto informace z pomocného souboru vloží ve formě odpovídajících zápisů do vysázeného dokumentu.

## Navržené řešení

Na vyhrazeném počítači je nainstalován překladač systému  $\text{T}_{\text{E}}\text{X}$  a další nezbytné podpůrné utility. Jde například o program `vlna`, který opatří zdrojový text, v místech, kde je třeba, nezlomitelnými mezerami nebo o program `dvips` generující z formátu DVI formát PostScript.

Celá instalace  $\text{T}_{\text{E}}\text{X}$ u a podpůrných programů je umístěna ve vyhrazeném prostoru, který je oddělen od vlastního operačního systému, především pak konfiguračních souborů a souborů obsahujících důležité informace, zneužitelné nepovolaným uživatelem. Odstínění je realizováno programem `chroot` a je tak zajištěna bezpečnost při nepovoleném vniknutí přes webové rozhraní.

Dalším prvkem bezpečnosti je omezení velikosti ukládaných souborů. V současné době je omezena na 1 MB. To by mělo postačovat pro vkládání obrázků, ale zároveň to ztěžuje útočnickovi zaplnit disk serveru a znemožnit tak fungování systému  $\text{T}_{\text{E}}\text{XonWeb}$ . Funkčnost vlastního serveru by ani toto nemělo ovlivnit neboť systémové soubory jsou umístěny na jiném logickém disku.

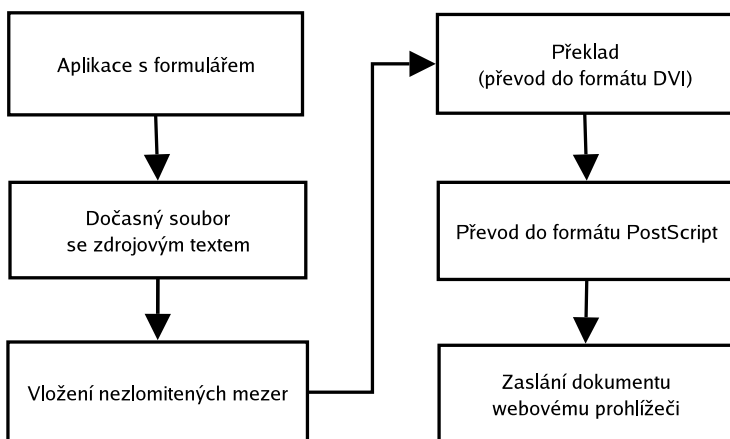
Princip fungování je v zásadě velice prostý. Využili jsme vlastnosti  $\text{T}_{\text{E}}\text{X}$ , že zpracovává data dávkovým způsobem, což nám umožňuje proces tvorby dokumentu algoritmizovat.

V okamžiku, kdy uživatel napíše zdrojový text dokumentu, připojí potřebné soubory, vybere typ překladače a stiskne tlačítko, odešlou se data na server. Tam je z textu obsaženého ve formuláři vytvořen dočasný soubor. Ten je opatřen příkazem `\nonstopmode` zajišťujícím neinteraktivnost překladu. Poté do něj program `vlna` dosadí na patřičná místa nezlomitelné mezery. Následně je soubor přeložen do formátu DVI a poté překonvertován do PostScriptu. V případě PDF je prováděn překlad rovnou do tohoto formátu. Výsledný tvar je poté zaslán zpět webovému klientovi, který jej zobrazí, popřípadě nabídne možnost uložení na lokální počítač. O průběhu překladu je vygenerován logový soubor, který si uživatel může prohlédnout a zjistit případné chyby. Pokud překlad nedopadne dobře a není možno výsledný tvar vygenerovat, systém o tom podá hlášení a nabídne možnost prohlédnutí logového souboru.

Jak bylo zmíněno výše, celý proces překladu probíhá na serveru ve vyhrazeném prostoru, aby se zamezilo případným nepovoleným akcím. Jde například o možnost připojení souboru `/etc/passwd` do dokumentu. Tento soubor obsahuje informace o uživatelských účtech na serveru. Takové informace by mohly případnému útočnickovi usnadnit průnik na server a provádění nekalých činností. Obrázek 1 schematicky naznačuje činnost systému.

## Uživatelské rozhraní

Pro přístup k systému lze použít jakýkoliv webový prohlížeč podporující formuláře a kaskádové styly. Doporučeny jsou zejména prohlížeče Mozilla a Internet Explorer. Základem stránky je formulář s textovým polem, do kterého uživatel zapisuje zdrojový text dokumentu a opatřuje jej značkami systému  $\text{T}_{\text{E}}\text{X}$ . Pod tímto oknem jsou tlačítka, která umožňují výstup buď ve formátu PostScript nebo PDF. Uživatel může roletkovým menu samozřejmě určit, který překladač se použije. Standardně jsou nabízeny  $\text{CSL}^{\text{A}}\text{T}_{\text{E}}\text{X}$ ,  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ ,  $\text{CSPlain}$ ,  $\text{T}_{\text{E}}\text{X}$  a to i ve variantách pro PDF. Počet průchodů překladače lze ovlivnit také, nabízeny jsou jeden, dva nebo tři. Jiné hodnoty v podstatě nemají význam. Další důležitou položkou ve formuláři je možnost připojovat vlastní soubory. Mohou to být jak obrázky, tak i jiné zdrojové soubory s textem dokumentu nebo soubory rozšiřujících maker. To umožňuje uživateli vytvořit strukturovaný dokument odděleně



**Obrázek 1.** Schéma činnosti navrženého systému „TeX on Web“

na vlastním počítači a poté jej pouze za použití webového rozhraní přeložit a získat potřebný výstupní formát.

## Pozadí systému

Pro tento účel byl vyhrazen jeden stroj architektury IBM PC s procesorem Celeron s nainstalovaným operačním systémem Linux (distribuce Fedora), webový server Apache, programovací jazyk Perl (mod\_perl) a typografický systém  $\text{\TeX}$  (distribuce  $\text{te\TeX}$ ).

Systém je v zásadě tvořen perlovským modulem, který obsahuje metody zajišťující veškerou práci s překladem různými překladači, převodem do nabízených formátů, ukládání vlastních souborů atd. Na něj je navázán perlovský skript vytvářející uživatelské rozhraní popsané výše.

## Dosavadní výsledky

V  $\text{\TeX}$ ové komunitě byl požadavek na takovýto projekt již několikrát vznesen, takže se dá usoudit, že o něj bude velký zájem. Důležitým faktem je, že usnadní seznamování začínajících uživatelů se systémem  $\text{\TeX}$ , případně nabídne možnost jednorázového použití lidem, kteří s  $\text{\TeX}$ em pracovat pravidelně nechtějí, ale jsou v situaci, kdy nemají na svém počítači  $\text{\TeX}$  nainstalován nebo se tomu z nějakého důvodu chtějí vyhnout. Jde třeba o případ tvorby diplomových prací, kdy student potřebuje provést kvalitní vysazení dokumentu a nechce se seznamovat s pozadím systému. Z tohoto důvodu je nabízena možnost použití balíku maker pro sazbu diplomových prací.

Popisované řešení ještě není kompletně implementováno. Jakmile však budou práce dokončeny, bude systém zveřejněn a vystaven důkladnému testování a připomínkování. Podnětné nápady budou samozřejmě zváženy a poté zapracovány.

Uvažujeme především o možnosti upravení prostoru pro vkládání zdrojových kódů (formulářový prvek `textfield`), který simuluje textový editor. Chceme aby se jednak pro větší přehlednost a snadnější odhalování chyb zvýrazňovala syntaxe vkládaných textů a aby se automaticky číslovaly řádky.

Dalším uvažovaným rozšířením je možnost uložení zdrojového textu na disk lokálního počítače. Dosavadní praxe, kdy uživatel vytvořený text označí do bloku a pak jej přes schránku uloží do souboru na svém počítači, se jeví jako poněkud těžkopádná.

Poslední problém, kterým se v současné době vážněji zabýváme, je dlouhá odezva systému na pomalejších linkách. Je to způsobeno především velkým objemem přenášených vysazených dokumentů. Tento problém se výrazněji projevuje u formátu PostScript. Nabízí se možnost výsledný dokument na serveru zkomprimovat a posílat jej v tomto tvaru, což sníží velikost přenášených dat. Komprimace se samozřejmě nebude provádět implicitně, ale jen v případě, že si to uživatel bude přát.

Budeme-li ovšem chtít tímto způsobem rozšiřovat možnosti nastavování systému, nevyhne se určitěmu zneprehlednění stránky. Stojí také za zamýšlení, zda nenastavovat všechny záležitosti samostatně v jiném formuláři, a na stránce, kde se vkládají data, mít jediné tlačítko pro vysazení dokumentu. Každá z variant má své výhody a nevýhody, pravděpodobně se optimální koncepce prosadí postupně podle požadavků uživatelů.

Testovací verze je přístupná na adrese <http://tex.mendelu.cz/>.

## Předdefinované styly

Zpřístupnění funkční instalace systému  $\text{T}_{\text{E}}\text{X}/\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  je motivováno zejména výukovými účely. Přejdem na webový systém můžeme upustit od komplikované správy instalací na různých počítačových učebnách.

Kromě tohoto prvního a zřejmého cíle se však ukazuje, že vzrůstá počet uživatelů, kteří sáhnou po zpracování dokumentů, k nimž existuje funkční a typograficky bezchybný sázečí styl. Tyto jednoznačně pozitivní zkušenosti již existují se stylem pro sazbu diplomových prací.

Je tedy přirozené pokusit se o vytvoření sady stylů pro dokumenty jiných typů. Může se jednat například o dopisy, objednávky, nejrůznější formuláře apod. Zde se otevírá široké pole působnosti, uvědomíme-li si, že takových dokumentů je na hlavním serveru univerzity vystaveno ve formátu Microsoft DOC asi 70. Tyto dokumenty jsou však připravovány zcela nahodile, neoborně a ve své sestavě bezkonceptně, což bohužel běžný uživatel naprosto nevnímá.

Vytvořením určité alternativy by bylo možné dokázat, že systém  $\text{T}_{\text{E}}\text{X}$  s jeho nadstavbami není jen hračkou pro vědce-matematiky.

Dalším stupněm zjednodušujícím využití standardních dokumentů je vytvoření webového rozhraní, kde jsou připravena pole pro vyplnění požadovaných textových segmentů, z nichž se automatizovaně sestaví, vysází a vrátí uživateli bezchybný dokument odpovídající normám a typografickým zásadám. Taková nadstavba je použitelná pro takovou třídu dokumentů, u nichž není počet textových segmentů příliš variabilní, ale takových dokumentů je ve zmíněné sestavě většina.

## Závěr

Projekt zpřístupňující precizní sazbu systémem  $\text{T}_{\text{E}}\text{X}$  včetně všech nastaveb je ve své počáteční fázi. Možnosti, které jsou nyní k dispozici, lze shrnout do těchto bodů:

- Zpřístupnění instalace začínajícímu uživateli nebo zájemci, který chce vidět základní možnosti.
- Snížení objemu údržby učebnových instalací systému  $\text{T}_{\text{E}}\text{X}$ .
- Využití připravených stylů pro tvorbu standardních dokumentů.
- Vytvoření další obalové aplikace využívající překladu jako jádra své činnosti a zpřístupňující standardní dokumenty přes webové rozhraní.

Z prvních ohlasů lze nabýt přesvědčení, že nabízené služby mohou být pro uživatele atraktivní a že se s minimálními nároky na údržbu může takový systém s úspěchem provozovat.

## Odkazy

1. ČSN 01 6910 – Úprava písemností psaných strojem nebo upravených textovými editory. Praha: Normalizační institut, 1997.
2. Seznam formulářů ke stažení. Dokument HTML s odkazy na soubory formátu Microsoft DOC. Dostupné na <http://www.mendelu.cz/formulare>. 20. května 2004.

---

---

## Ligatura aneb začínáme s $\text{T}_{\text{E}}\text{X}$ em

MILAN ŠORM

Z dlouholetých zkušeností s  $\text{T}_{\text{E}}\text{X}$ em pozoruji, že nainstalovat nějakou  $\text{T}_{\text{E}}\text{X}$ ovou distribuci na systém Windows s fungující češtinou, editorem, zobrazováním a příp. připravenou nápovědou či ukázkovými styly je pro