

# Postup převodu studijních materiálů do formátu EED

Verze I

Středisko ELSA ČVUT v Praze

Wanda Gonzúrová, Pavel Hrabák

Poslední aktualizace 1. dubna 2015

Tento text vznikl za podpory projektu ESF OPVK ExpIn, *Sít' expertních pracovišť k zajištění inkluze v terciárním vzdělávání*, reg. č. CZ.1.07/2.2.00/29.0010.





## Obsah

Úvodem .....	4
Přehled postupu úpravy textů do formátu EED .....	5
1. Fáze převodu pomocí OCR .....	6
2. Fáze formátování textu .....	8
3. Fáze přepisu matematiky .....	10
4. Fáze kompletace .....	10
Reference .....	11

## Úvodem

Tento dokument vznikl za účelem sjednocení a optimalizace pracovního postupu adaptace studijních materiálů prováděné ve středisku ELSA. Pracovní postup zahrnuje nejen popis jednotlivých fází a doporučená nastavení používaného softwaru, ale i přesný popis dílčích výstupů. Nutnost takto podrobného a do jednotlivých kroků rozpracovaného návodu je dána potřebou sjednotit práci většího počtu externích digitalizátorů, kteří se rekrutují z řad studentů ČVUT. Dodržováním tohoto postupu je zaručena jednotná podoba výsledných dokumentů navzdory různorodosti a proměnlivosti skupiny externích digitalizátorů.

Níže uvedený návod vede externího digitalizátora postupem adaptace krok za krokem. Konkrétní pravidla formátování a zápisu matematiky jsou uvedena v [1], [2], vycházející z [3].

Vzhledem k vývoji kompenzačních pomůcek a studijních strategií je tento dokument považován za neuzavřený a bude dále inovován dle relevantních připomínek digitalizátorů a uživatelů adaptovaných materiálů.

Wanda Gonzúrová a Pavel Hrabák

## Přehled postupu úpravy textů do formátu EED

Převod titulu do formátu EED probíhá ve čtyřech fázích:

1. **Převod pomocí OCR,**
2. **Formátování textu,**
3. **Přepis matematiky,**
4. **Kompletace.**

Je nutné, aby digitalizátor zvládal první tři fáze v plné míře, i kdyby se podílel jen na jedné z nich. Fázi kompletace provádí kmenový pracovník střediska.

Konečné výstupy po provedení všech fází jsou:

- **NázevTitulu** - soubor ve formátu Fine Reader
- **NázevTitulu\_ZDROJ.pdf** - soubor obsahující obrazy skenu, případně zdrojový soubor ve formátu PDF, byl-li k dispozici
- **NázevTitulu\_OCR.doc** - neopravený dokument bez formátování (nearchivuje se)
- **NázevTitulu\_FORM.doc** - opravený formátovaný dokument, matematické výrazy označeny \$\$
- **NázevTitulu\_FINAL.doc** - kompletně převedený dokument včetně matematických výrazů
- **Příloha** - složka obsahující obrázky ve formátu PNG, JPG, či PDF.
- **NázevTitulu.doc** - finální dokument zkompletovaný kmenovým pracovníkem

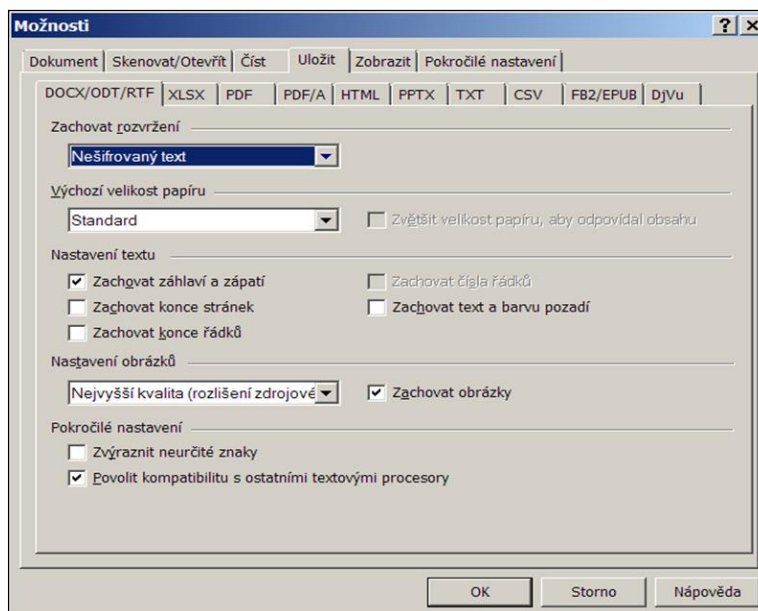
Soubory **NázevTitulu**, **NázevTitulu\_ZDROJ.pdf** a **NázevTitulu\_FORM.doc** se archivují pro potřeby úprav do jiného formátu EED.

# 1. Fáze převodu pomocí OCR

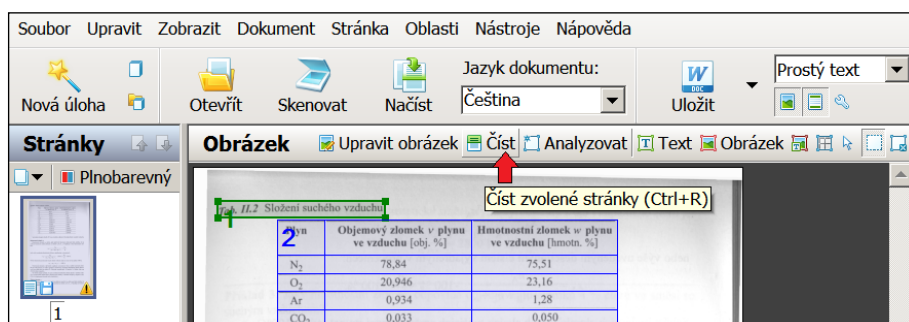
V první fázi je nutné převést předlohu do textového formátu. Tištěnou předlohu je nutné naskenovat (a to včetně tiráže!), elektronickou převést do formátu kompatibilního s použitým OCR softwarem (např. PDF, TIFF). Poskytne-li autor zdrojový soubor titulu (LaTeX, Word, apod.), fáze OCR odpadá a v dalších fázích je třeba postupovat individuálně dle formátu zdrojového dokumentu.

Doporučený postup pro rozpoznání textu pomocí FineReader, verze 11 (dále FR)

1. Nastavit: Nástroje/Možnosti/Uložit/DOCX-ODT-RTF/Zachovat rozvržení: Nešifrovaný text a Nastavení obrázků: Zachovat obrázky



2. FR automaticky analyzuje obrazovou předlohu. Před uložením rozpoznaného textu je třeba zkontrolovat a upravit analyzované oblasti textu a obrázků. Po úpravě oblastí je nutné nechat dokument znovu „čist“.



3. Po rozpoznání uložit dokument aplikace FineReader (obsahuje obrázky a informace o posledním rozpoznání).
4. Obrázky uložit do vícestránkového souboru PDF: Soubor/Uložit dokument jako/Dokument PDF/A, zvolit Vytvořit samostatný soubor pro všechny stránky. Takto uložený soubor obsahuje kromě obrázků i textovou vrstvu z rozpoznání.
5. Rozpoznaný dokument uložit jako DOC.
6. V rozpoznaném souboru DOC vymazat veškeré formátování (Změnit Styly/Vymazat vše) a případné hypertextové odkazy (Ctrl+A, Ctrl+Shift+F9). Uložit.
7. Otevřít šablonu: EED-šablona.dotx/.dotm/.dot.
8. Vložit text z rozpoznaného souboru prostřednictvím karty Vložení/Objekt/Text ze souboru/rozpoznaný soubor DOC.
9. Uložit jako **NázevTitulu\_OCR.doc**.

Výstup:

- **NázevTitulu\_OCR.doc** bez formátování.
- Složka **Příloha** obsahující obrázky ve formátu JPG, PNG či PDF.<sup>1</sup> Pojmenování obrazových souborů viz [1], Kap. 4.1.
- Soubor **NázevTitulu** ve formátu Fine Reader
- Soubor obrázků předlohy **NázevTitulu\_ZDROJ.pdf**

---

<sup>1</sup> Lze uložit všechny obrazové soubory najednou: Uložit soubor **NázevTitulu\_OCR** jako DOCX, otevřít pomocí WinRar (nebo podobného programu), obrázky jsou ve složce word/media.

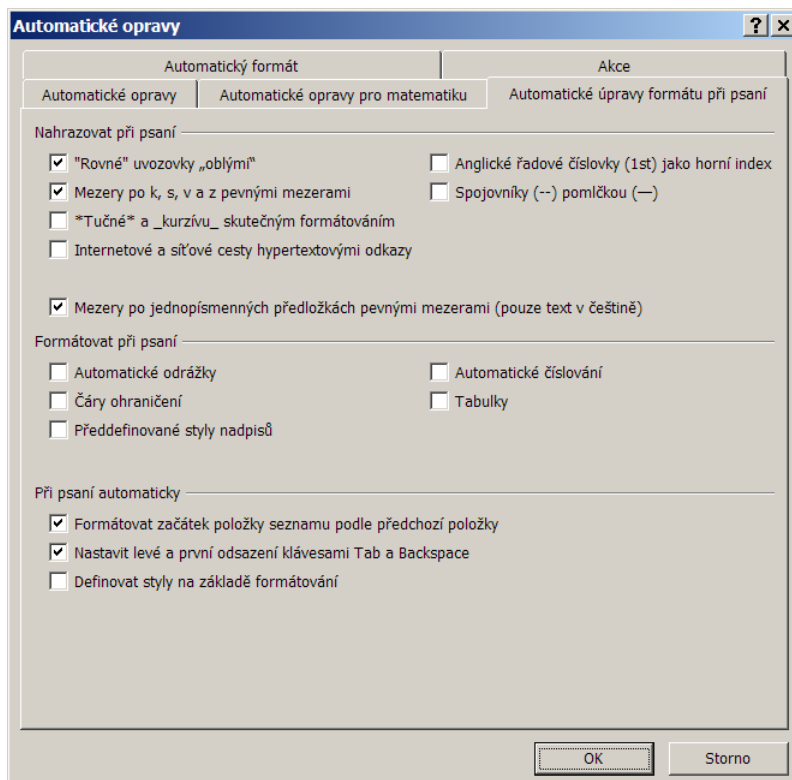
## 2. Fáze formátování textu

Ve druhé fázi se doplňuje struktura a formát dokumentu a opravují se chyby v rozpoznání dle *Pravidel pro převod studijních materiálů do formátu EED I a EED II*, viz [1], u jednotlivých bodů jsou specifikovány příslušné kapitoly:

- Oprava chyb v rozpoznání a překlepů (Kap. 2).
- Formátování textu, obrázků a tabulek (Kap. 3 a 4).
- Struktura rovnic na samostatném řádku (Kap. 5).
- Prostor pro zápis matematických výrazů je označen dvěma dolary \$\$.
- Nestandardní úpravy nebo úpravy, které jsou v [1] označeny jako „průvodky-hodné“, se průběžně zaznamenávají do kapitoly **Průvodka** na začátku dokumentu.

Doporučené nastavení MS Word:

- Zakázat automatické opravy i automatické opravy pro matematiku.
- V záložkách Automatické formátování a Automatické úpravy formátu při psaní v kolonce Nahradit při psaní zakázat vše kromě Nahradit rovné uvozovky oblými a Mezery po k, s, v a z pevnými mezerami:





Výsledný dokument této fáze obsahuje:

- Text stylem Normální, Nadpisy úrovně 1, 2, 3, ..., seznam a poznámky pod čarou.
- Navigační symboly před nadpisy, tabulky, obrázky, čísla rovnic, apod. (#, @, \*, &).
- Doplněné řezy písma.
- Zpracované obrázky, grafy a tabulky včetně příslušných hypertextových odkazů.
- Prostor pro matematické výrazy označený dvěma dolary \$\$.
- Kapitulu Průvodka na začátku dokumentu.
- Hypertextový obsah za průvodkou

Vstup:

- **NázevTitulu\_OCR.doc** bez formátování a hypertextových odkazů.
- Složka **Příloha** s obrázky ve formátu JPG, PNG, PDF a pod.

Výstup:

- Soubor **NázevTitulu\_FORM.doc**, opravený a formátovaný.
- Složka **Příloha** obsahující obrázky, na které vedou hypertextové odkazy.

### 3. Fáze přepisu matematiky

Do (nebo místo) připravených dolarů \$\$ se vpisují matematické výrazy dle odpovídajícího formátu EED podle [1] Kap. 5.

- **EED I:** Veškeré matematické výrazy včetně jednopísmenných znaků se přepisují dle pravidel Blind Friendly LaTeX [2] (dále BFL) a nepřevádějí se do grafické podoby.
- **EED II:** Matematické výrazy se zapisují dle BFL nebo přímo z klávesnice. Výrazy psané pomocí BFL jsou převedené do grafické podoby pomocí MathType.
- Nestandardní úpravy, nebo úpravy, které jsou v [1] či [2] označeny jako „průvodkyhodné“, se průběžně zaznamenávají do kapitoly **Průvodka** na začátku dokumentu.

Vstup:

- Soubor **NázevTitulu\_FORM.doc**, opravený a formátovaný.
- Složka **Příloha** obsahující obrázky, na které vedou hypertextové odkazy.

Výstup:

- Soubor **NázevTitulu\_FINAL.doc** opatřený průvodkou s poznámkami k formátování a zápisu matematiky.
- Složka **Příloha** obsahující obrázky, na které vedou hypertextové odkazy.

### 4. Fáze kompletace

Kompletaci provádí kmenový pracovník střediska ELSA. Dokument je opatřen řádnou průvodkou, tiráží a dalšími informacemi dle dispozic. Příslušné dokumenty jsou archivovány a katalogizovány.

## Reference

- [1] W. Gonzúrová a P. Hrabák, *Pravidla pro převod studijních materiálů do formátu EED I a EED II*, Praha: Středisko ELSA, ČVUT, 2015.  
Dostupné online: [http://www.elsa.cvut.cz/\\_media/files/EED\\_Verze\\_I.pdf](http://www.elsa.cvut.cz/_media/files/EED_Verze_I.pdf)
- [2] W. Gonzúrová a P. Hrabák, *Příručka pro přepis matematického textu pomocí Blind Friendly LaTeX (BFL)*, Praha: Středisko ELSA, ČVUT, 2015.  
Dostupné online: [http://www.elsa.cvut.cz/\\_media/files/BFL\\_Verze\\_I.pdf](http://www.elsa.cvut.cz/_media/files/BFL_Verze_I.pdf)
- [3] M. Hanousková, *Metodika k úpravám elektronických textů pro zrakově postižené uživatele, Verze VII*, Brno: Středisko Teiresiás, Masarykova univerzita, 2010.  
Dostupné online: [http://www.teiresias.muni.cz/download/Metodika\\_VII.pdf](http://www.teiresias.muni.cz/download/Metodika_VII.pdf)