

SKENOVÁNÍ, OCR A FORMÁT PDF

Co jste kdy chtěli vědět a nezeptali se...

(lekce 3)

Karel Hrubeš

Břetislav Verner

Obsah a redakce textu: Břetislav Verner

© Autoři lekce, ver. 2023: Tento kurz je určen pro výuku zrakově postižených uživatelů speciálního zvětšovacího a odcítacího programu ZoomText nebo Fusion a autoři kurzu nabízejí pro tento účel celý text nebo jeho části k bezplatnému použití a šíření. Zároveň však autoři prosí, aby přitom vždy byl uváděn odkaz na tento původní zdroj, z něhož použití textu kurzu nebo jeho částí pochází (viz „licenci“ v poslední kapitole).

Názvy produktů, firem apod. použité v textu mohou být ochrannými známkami nebo registrovanými ochrannými známkami příslušných vlastníků.

OBSAH

1	Prolog.....	3
2	Co je na tom papíru napsáno? Digitalizovat !	3
2.1	Víte, že kamerová lupa je v ZoomTextu?.....	3
2.2	Digitalizovat a číst hlasem	3
3	Co je OCR a technologie rozpoznávání?	4
4	Jaké jsou prostředky pro OCR?	5
4.1	Webová služba	5
4.2	Počítačová aplikace	6
5	Rozpoznávací program ABBYY FineReader PDF.....	7
5.1	Spuštění, uživatelské rozhraní a ovládání programu	7
5.2	Nastavení programu: dialog Možnosti.....	9
5.3	Nová úloha - Otevřít.....	9
5.4	Nová úloha - Skenovat.....	11
5.5	Nová úloha – Porovnat	13
5.6	Systémová integrace FineReaderu.....	13
6	Formát PDF	14
6.1	Typy formátů PDF.....	15
6.2	Asistivní čtečka PDF v prohlížeči Edge.....	16
6.3	Jak tedy přečíst nepřístupné PDF? Úkol pro OCR !	16
7	Bonusy	18
7.1	Malý test přístupnosti a nepřístupnosti PDF	18
7.2	Jak lze získat strukturované PDF	19
7.3	Jak převést dokument PDF do formátu Wordu?	19
7.4	Jak nouzově na nepřístupná PDF.....	19
7.5	Rozpoznání obrazovky: ScreenShot Reader	20
8	Některé klávesové zkratky z této lekce	20
9	Jak dále	21
9.1	Licence	22
9.2	Poděkování a kontakt	22

1 Prolog

V této výukové lekci si řekneme, co znamená digitalizace, co je to tajemné OCR čili rozpoznávání textu a jak si hlasovým výstupem přečtete, co je napsáno na papírové předloze. V průběhu lekce se také seznámíte s programem FineReader a s tím, jak ho mohou využívat zrakově postižení uživatelé, kteří pracují se speciálními programy ZoomText, Fusion či Jaws. Nakonec se budeme věnovat formátu PDF, jeho přednostem i nechtostem a jak jej zpřístupnit.

2 Co je na tom papíru napsáno? Digitalizovat !

Pokud ještě můžete číst očima, jenom potřebujete hodně velké zvětšení, nemusíte nic digitalizovat, stačí papírový dokument vložit pod kamerovou lupu a číst.

2.1 Víte, že kamerová lupa je v ZoomTextu?

Nástroj Kamera v ZoomTextu vám umožní použít běžnou webovou kameru s vysokým rozlišením (HD) k zobrazování tištěných materiálů a dalších předmětů na obrazovce počítače. Obraz můžete zvětšit či zmenšit, upravit si jeho kontrast. Můžete také přepínat mezi celoobrazovkovým zobrazením a dělenou obrazovkou, která dovoluje souběžně sledovat obraz z kamery a prostředí Windows.

Kamerovou lupu najdete v rozhraní ZoomTextu na kartě Nástroje pod ikonou Kamera. Zde ji nastavíte a spustíte (zkratka CapsLock + Ctrl + C).

2.2 Digitalizovat a číst hlasem

Pokud si však chcete informace z papíru poslechnout hlasovým výstupem, číst jej na braillovém řádku nebo s textem papírového dokumentu dále pracovat, například ho upravovat, někomu dalšímu poslat apod., musíte dokument digitalizovat. Digitalizovat znamená předlohu převést do počítače, do počítačem zpracovatelné podoby.

Dejme tomu, že chcete digitalizovat článek z časopisu nebo vytištěnou stránku. Kdybyste digitalizovali tak, že byste text článku přepisovali na klávesnici do počítače, mohli byste tím strávit hodiny či dokonce dny. Mnohem rychlejší a pohodlnější je automatizovaná digitalizace pomocí skeneru nebo digitální kamery.

Skener nebo kamera sama nestačí k tomu, abyste textovou informaci z papíru mohli editovat (upravovat) třeba ve Wordu nebo si ji přečíst hlasovým výstupem. Všechno, co skener či kamera udělá, je, že vytvoří digitální fotografii dokumentu (tzv. rastrový obraz). Abyste dokázali využít či syntetickým hlasem přečíst data z naskenovaného dokumentu (z toho bodového rastru), potřebujete počítačový program OCR, který z naskenovaného obrázku vytáhne jednotlivá písmena, složí je do slov, do vět a do celého textu a tím vám textovou informaci z papírového dokumentu zpřístupní a umožní ji editovat nebo si ji poslechnout.

Digitalizace s rozpoznáním textu tedy obnáší tři kroky:

- Sejmout předlohu (skenerem, digitální kamerou)
- Rozpoznat (OCR programem)
- Uložit rozpoznáný dokument (jako soubor do počítače).

3 Co je OCR a technologie rozpoznávání?

OCR je zkratka anglického termínu Optical Character Recognition (česky „optické rozpoznání znaků“) a funguje tak, že z naskenovaných dat „vytáhne“ text a jeho jednotlivá písmena porovnává se vzorovými sadami různých druhů písma. Rozpoznávání se ještě zpřesňuje používáním slovníků.

Obecně lze říci, že OCR je technologie, která převede text vyskytující se v obrázcích sejmutých skenerem nebo digitální kamerou do editovatelné a pro odečítače či hlasové výstupy čitelné podoby. S rozpoznáním textem pak můžete běžně pracovat v textových editorech (jako Word, Poznámkový blok apod.) a v rozpoznaném textu vyhledávat slova či sousloví (fráze).

Chyby v rozpoznávání

Rozpoznání písma nemusí být bez chyb, vady tisku, nekvalitně sejmutá či naskenovaná předloha obvykle vedou k větší chybovosti při rozpoznávání textu programy OCR – s tím je nutné počítat, byť kvalita rozpoznávání dobrých OCR programů jako je například FineReader, je velmi vysoká. Chyby je však potom třeba pro některé účely použití opravit.

Primárním zdrojem chyb může být sama předloha, když obsahuje vady tisku, například slitá písmena nebo slitou diakritiku. Zdrojem chyb také může být příliš malé písmo, resp. malé rozlišení při skenování nebo fotografování. Také se může stát, je-li papír předlohy velmi tenký, že prosvítá písmo nebo grafika z opačné strany skenovaného listu, která se pak na digitalizované stránce smíchá s textem lícové strany. Problémem také může být velmi členitý text (text rozházený mezi grafikou) nebo text členěný do složitější struktury sloupců (s jednoduchou strukturou sloupců svázaných jedním nadpisem se lepší OCR programy dokážou vyrovnat).

Další chyby mohou vznikat v samotném procesu digitalizace, jednu jsme již naznačili – špatně nastavené rozlišení.

Při skenování by nastavené rozlišení pro skenování standardního písma (časopis, kniha) mělo být 250 až 300 dpi, při drobném písmu, například pro text příbalového letáku léků, pak vyšší třeba až 400-500 dpi (dpi je jednotkou rozlišení a znamená počet obrazových bodů na délkový palec). Dalším problémem při skenování může být našikmo založená stránka či silnější kniha a ohyb resp. nedostupnost textu blízko hřbetu knihy: takový text může být nekvalitně naskenovaný a v důsledku pak chybně rozpoznáný.

Fotografování digitální kamerou není úplně ekvivalentní skenování, mohou zde vznikat chyby i z jiných příčin. Některé zdroje chyb jsou společné, jiné jsou spíše typické pro kameru.

Podobným zdrojem chyb bývá malé rozlišení (měli byste dbát, abyste fotografovali alespoň v rozlišení HD) nebo šikmo umístěný text pod kamerou.

Specifickým zdrojem chyb při snímání kamerou může být

- Málo osvětlená nebo nerovnoměrně osvětlená předloha: v tmavých částech obrazu pak bývá nerozpoznaný nebo chybně rozpoznáný text.
- Nekvalitně vyfotografovaný text na okrajích snímku: málo osvětlený, prohnutá stránka blízko hřbetu.
- Případně špatně nastavený záběr kamery, který rozřízne slovo – neúplná slova či neúplné texty nebývají dobře rozpoznané.

Shrneme-li, lze říci, že nejčastějšími chybami při rozpoznávání textu jsou: chybně rozpoznané znaky (slité, blízko sebe či speciální), chybné uspořádání textu do sloupců, chybné řazení textu v tabulkách, chybně rozpoznáný jazyk při střídání různých jazyků v jednom textu, nerozpoznané formátování (atributy fontů, text oddělený do sloupců, do záhlaví či zápatí).

A platí: Čím lepší OCR program, tím menší chybovost a přesněji rozpoznané formátování, a čím lépe rozpoznané formátování, tím blíže je rozpoznáný dokument svému originálu.

4 Jaké jsou prostředky pro OCR?

Jsou v podstatě dvě možnosti, jak rozpoznat text v naskenovaném (vyfotografovaném) dokumentu:

- Webovou službou, k níž potřebujete funkční internetové připojení
- Počítačovým OCR programem, k němu internet většinou nepotřebujete

4.1 Webová služba

Připojíte se k serveru, který takovou službu poskytuje, na server nahrajete soubor ze skeneru nebo z digitální kamery a spustíte převod do požadovaného formátu s rozpoznáním textem (obvykle do Wordu). Za chvíli (záleží na velikosti zasláního dokumentu) dostanete nabídku, abyste si rozpoznáný soubor stáhli do počítače – a je hotovo.

Podobných služeb je mnoho, některé jsou zdarma, některé jsou placené, ty zdarma obvykle mívají některá omezení na velikost souborů k rozpoznávání či omezený počet souborů k rozpoznávání za nějakou dobu (den, měsíc).

Službu vyhledáte například tak, že do vyhledávače zadáte frázi „ocr online“. Mějte však na paměti, že online nástroje nemusí zaručit bezpečnost textu a soukromí a buďte obezřetní při jejich využívání.

4.2 Počítačová aplikace

Rozpoznávacích počítačových programů existuje mnoho, některé jsou zdarma (freeware), jiné jsou placené.

Freewarové programy OCR obvykle formátované rozpoznávání nenabízejí, nabízejí jen převod (rozpoznání) do prostého textu. To mnohdy stačí, když například s textem nechcete dále pracovat a chcete si jen poslechnout, o čem ten dokument je.

Velmi kvalitním rozpoznávacím programem (placeným) je program FineReader, kterému se budeme v následující kapitole věnovat. Tento program je do velké míry zpřístupněný, takže s ním mohou pracovat i nevidomí uživatelé vybavení vhodným odečítacím programem (např. Fusion, Jaws).

Dobré zpřístupnění může být úskalím některých freewarových OCR programů, při volbě vhodného programu pro rozpoznávání textu je tedy nutné dávat pozor i na toto.

Fusion/Jaws: OCR přímo do Wordu

Technologie OCR zabudovaná v těchto odečítačích zpřístupňuje text v grafice a rozpoznávaný text pošle přímo do Wordu. Pokud jste si nainstalovali kancelářský balík Microsoft Office, můžete:

- Příkazem INSERT+MEZERNÍK, pak O a pak R spustit OCR na zvoleném obrázku v Průzkumníku souborů nebo na ploše a rozpoznávaný text pak otevřete přímo ve Wordu. Alternativně můžete také stisknout klávesu KONTEXTOVÉ MENU a v menu pak vybrat „Rozpoznávání textu do Wordu s Jaws“.
- Použijte odkaz „Otevřít v aplikaci Word“, který se po ukončení OCR procesu nachází dole v okně „Výsledky rozpoznávání textu“.

Nebo

- V Průzkumníku souborů kombinací Shift + F10 otevřete kontextové menu na zvoleném obrázku, jehož text chcete rozpoznávat. Zde pak v nabídce vybíráte ze dvou položek "Rozpoznávání textu do Wordu s Jaws" a Rozpoznávání textu s Jaws.

Poznámka: Zahájené rozpoznávání ukončíte funkcí „Zrušit rozpoznávání v průběhu“ příkazem INSERT + MEZERNÍK, pak O, pak Q. Náповědu k OCR získáte po zadání příkazu INSERT+MEZERNÍK, pak O, pak PODTRŽENÍ (čili otazník na anglické klávesnici).

5 Rozpoznávací program ABBYY FineReader PDF

FineReader je velmi kvalitní OCR program určený i pro náročnější práci – velmi dobře rozpoznává formátování původního textu a při rozpoznávání vykazuje malou chybovost.

V této kapitole si ukážeme, jak se pracuje s OCR programem FineReader verze 14, 15 či 16 (ovládání programu v té které verzi se může drobně lišit).

5.1 Spuštění, uživatelské rozhraní a ovládání programu

Při práci s programem FineReader postupujete v těchto krocích

- Spustit program
- Pořídit či otevřít dokument (skenovat, vyfotografovat nebo z uloženého souboru)
- Rozpoznat dokument
- Uložit rozpoznáný dokument ve vhodném formátu (.DOC, .RTF, .XLS, .PDF, .HTML, .TXT etc.) nebo přímo otevřít v cílové aplikaci (Word, Excel či Adobe Acrobat)

Bohužel ne všechny nabízené úlohy jsou dobře zpřístupněné pro odečítače obrazovky. Dále se budeme věnovat především těm funkcím FineReaderu, které zpřístupněné jsou.

5.1.1 Spustit program

Program ABBYY FineReader PDF lze spustit více způsoby:

- Odklepnutím ikony FineReader na ploše
- Prostřednictvím zástupce ABBYY FineReader v nabídce Start
- Napsáním části řetězce názvu programu do vyhledávacího pole Windows (stisknete klávesu WIN a hned začnete psát)

5.1.2 Volba úlohy

Po spuštění programu se ocitnete v jeho hlavním okně s názvem „Nová úloha“, které vlevo pod hlavním menu nabízí seznam těchto skupin funkcí

- **Otevřít:** Zde otevřete existující uložený dokument, který převedete do jiného formátu přístupného pro asistivní programy (odečítače). Skupina obsahuje funkce Otevřít dokument PDF, Otevřít v editoru OCR, Převést do PDF, Převést do MS Word, Převést do MS Excel, Převést do jiných formátů.
- **Skenovat:** Zde dokument digitalizujete skenováním a převedete do dalších formátů podobně jako ve skupině první. Skupina obsahuje funkce Skenovat do editoru OCR, Skenovat do souboru PDF, Skenovat do MS Word, Skenovat do MS Excel, Skenovat do obrazových souborů, Skenovat do jiných formátů.

- **Porovnat:** Funkce porovnává dva dokumenty. Funkce je bohužel nepřístupná pro odečítače obrazovky, k jejímu použití je nutné ovládání myši a vizuální kontrola.
- **Nedávné:** Nabídne seznam naposledy zpracovávaných dokumentů.
- **Mobilní aplikace:** V případě aplikace instalované v chytrém mobilu iOS nebo Android za použití cloudového úložiště Google Drive lze na mobilu pořízené dokumenty zpracovávat i v programu instalovaném v PC.

Začnete tím, že si ze seznamu vyberete úlohu, kterou chcete provést. Po seznamu se pohybujete ŠIPKOU DOLŮ/NAHORU, funkce jednotlivých skupin pak rozbalíte MEZERNÍKEM nebo klávesou ENTER.

Dále se budeme věnovat především prvním dvěma úlohám Otevřít a Skenovat, které asi budete používat nejčastěji.

5.1.3 Hlavní menu a panel nástrojů

Hlavní okno programu má také aplikační menu, do kterého vstoupíte obvyklým stiskem levého ALT nebo otevřete kombinací ALT+D. Po menu se pohybujete obvyklým způsobem a můžete v něm aktivovat dostupnou funkci. Menu opustíte buď aktivací některé funkce nebo bez vykonání akce stiskem ESC nebo opětovným stiskem levého ALT.

Pod hlavním menu je pruh nástrojů s panelem rychlého přístupu, které můžete potlačit z menu Zobrazení > Panely nástrojů > a zaškrtnout nebo odškrtnout další dvě samostatné položky „Hlavní panel nástrojů“, „Panel rychlého přístupu“. Po odškrtnutí obou položek obsahuje hlavní okno programu jen hlavní menu a pracovní prostor, což může vyhovovat uživatelům odečítacího programu.

V hlavním menu jsou dostupné tyto nabídky:

- **Soubor:** Příkazy pro otevření či pořízení dokumentu, pro převody dokumentů
- **Upravit:** Navigační příkazy a příkazy pro výběr či odstranění
- **Zobrazit:** Nabídka úloh paralelní k oknu Nová úloha
- **Nástroje:** Některé další úlohy a nástroje
- **Nápověda:** Odkaz na stránku s nápovědou, informace o aplikaci

Z hlavního menu je také dostupný dialog **Možnosti**, dialog pro nastavení programu (z menu Soubor > Možnosti nebo Nástroje > Možnosti).

5.1.4 Ovládání

Po spuštění patřičné funkce programu se ocitnete v prostředí dialogových oken, která ovládáte stejně jako v jiných přístupných aplikacích – dialogy procházíte tabulátorem, šipkami přepínáte přepínače, mezerníkem měníte zaškrtnutí nebo aktivujete tlačítko, vyplňujete editační pole, atd.

5.2 Nastavení programu: dialog Možnosti

Základní nastavení programu najdete v dialogu Možnosti, který otevřete z hlavního menu volbou Soubor > Možnosti, v některých jiných oknech volbou Nástroje > Možnosti. Pro nastavení programu je k dispozici šest karet v dialogu Možnosti: Obecné, Zpracování obrázků, Jazyky, OCR, Nastavení formátu, Jiné.

Mezi kartami přecházíte šipkami doprava a doleva. Pokud jste na názvu některé z karet, pohybujete se po jednotlivých možnostech otevřené karty tabulátorem. Změnu nastavení potvrdíte tlačítkem OK. Na kartách nastavujete používaný skener, čtečku PDF Reader, automatické úpravy při skenování, výběr jazyka pro rozpoznávání textu, také změnu jazyka uživatelského rozhraní a další nastavení.

Možných nastavení je velmi mnoho, některá jsou snadno pochopitelná, k jiným jsou potřebné hlubší znalosti, často také záleží na preferencích konkrétního uživatele.

Nemá smysl, abychom v této lekci probírali a vykládali různé možnosti nastavení programu – to najdete v nápovědě. Obecně lze pro začátek doporučit „**ponechte předvolená nastavení**“ a jestliže s různými nastaveními provádíte pokusy, pamatujte, že se vždy dokážete vrátit k těm předvoleným tlačítkem „Obnovit výchozí“ na kartě Jiné v dialogu Možnosti.

Individuálně však v každém případě musíte nastavit skener a jazyk OCR:

- Vybrat skener (Možnosti > Obecné > Zvolit zařízení): pod tlačítkem „Zvolit zařízení“ je seznam instalovaných skenerů, ze seznamu jeden vyberete.
- Vybrat jazyky OCR (Možnosti > Jazyky): pod tlačítkem „Zadat“ zaškrtnete v seznamu požadované jazyky pro automatický výběr nebo jen češtinu pro výběr ručně.

5.3 Nová úloha - Otevřít

Ke zpracování existujících souborů a práci s nimi slouží funkce obsažené ve skupině **Otevřít**. V této skupině úloh naleznete funkce, které umožňují pracovat s novými dokumenty nebo s rozpracovanými dokumenty uloženými v počítači nebo jiném dostupném úložišti.

Začneme druhou funkcí ve skupině, funkcí Otevřít dokument PDF se budeme zabývat až v následující kapitole.

5.3.1 Otevřít v editoru OCR

Tato funkce je určena pro zpracování uložených naskenovaných obrázků v grafických formátech nebo pro přidání dalších skenovaných stránek k rozpracovanému dokumentu.

Po aktivaci funkce se zobrazí standardní dialogové okno pro otevření existujícího souboru (ekvivalentně také Soubor > Otevřít v editoru OCR).

Výběrem existujícího grafického souboru (obrázku) se soubor otevře v editoru OCR, který má přednastavena dvě podokna: v levém podokně Obrázek se otevře vybraný grafický soubor, v pravém podokně Text se po chvíli objeví rozpoznaný text z originálu vlevo.

Obecně jsou v pracovním prostoru editoru OCR čtyři podokna

Stránky (ve svislém pruhu vlevo): Obsahuje miniatury jednotlivých stran dokumentu, z nich vybíráte zobrazovanou stránku. Podokno skryjete z menu Zobrazit > Zobrazení okna Stránky > Zobrazit okno Stránky (F5).

Obraz (v pracovním prostoru vlevo): Ukazuje grafický obraz aktuální stránky. Zde lze upravovat oblasti rozpoznávání a měnit vlastnosti oblastí. Je to okno čistě grafické, které odečítač nepřečte. Kdo však pracuje na obrazovce očima, tak se může pomocí ZoomTextu v tomto okně pohybovat a oblasti upravovat. Podokno skryjete z menu Zobrazit > Okna Obraz a Text > Zobrazit jen okno Text (F8).

Text (v pracovním prostoru vpravo): To je textové okno s rozpoznáním textem otevřeného souboru, text je odečítačem čitelný a text lze zde upravovat. Podokno skryjete z menu Zobrazit > Okna Obraz a Text, nicméně skrýt okno Text asi smysl nedává.

Lupa (v pracovním prostoru dole): Zobrazuje zvětšený obraz aktuálního řádku textu, nebo jakékoli jiné oblasti obrazu, kde aktuálně pracujete. V podokně Obrázek vidíte celou stránku a v podokně Lupa můžete podrobněji studovat určité části obrazu, např. porovnávat obraz znaku s jeho rozpoznáním protějškem v okně Text. Podokno je grafické a tedy pro odečítače nepřístupné. Podokno skryjete z menu Zobrazit > Okno Lupa > Zobrazit okno Lupa (CTRL + F5). Asi je rozumné podokno Lupa schovat, pracovní prostor je pak přehlednější.

Chcete-li obrázek či text v podokně zvětšit nebo zmenšit, můžete pro to použít ovládací prvky ve spodní části podokna nebo klávesové zkratky Ctrl + Plus, Ctrl + Minus, které se aplikují na aktivní podokno.

K další práci se souborem lze použít různé funkce, které najdete v hlavní nabídce okna editoru nebo v jeho pásu příkazů. Hlavní nabídku lze otevřít pomocí jedné z klávesových zkratk přiřazených jednotlivým skupinám. Tuto klávesovou zkratku odečítač sdělí při přechodu mezi skupinami, například Soubor ALT+S, Upravit ALT+U, Zobrazit ALT+O, Rozpoznat ALT+R, Oblast ALT+A, Nástroje ALT+T, Nápověda ALT+N.

Výsledek práce je vhodné si uložit nebo exportovat do některého z nabízených formátů. K tomu slouží funkce Uložit jako a Odeslat do.

5.3.2 Převést dokument

Následující čtyři funkce pro převody dokumentů mají stejné rozhraní, proto je uvádíme najednou.

Poznámka: Převody dokumentů popisujeme zde, nicméně díky systémové integraci FineReaderu (viz dále podkapitolu „Systémová integrace FineReaderu“) jsou snadněji dostupné v kontextovém menu Průzkumníku nebo v jiných správčích souborů.

K dispozici jsou tyto funkce pro převody dokumentů:

- Převést do formátu PDF
- Převést do aplikace Microsoft Word
- Převést do aplikace Microsoft Excel
- Převést do jiného formátu.

Každá z těchto funkcí je ve skupině Otevřít reprezentována samostatným tlačítkem, které najdete pomocí šipek. Chcete-li spustit požadovanou funkci, stiskněte klávesu ENTER na příslušném tlačítku.

Po spuštění funkce se zobrazí standardní dialogové okno pro vyhledání a otevření souboru z dostupného úložiště. Po výběru souboru se zobrazí dialog Možnosti s nastavením parametrů převodu. Jednotlivé možnosti procházíte tabulátorem a převod potvrdíte tlačítkem Převést. V dalším kroku se aplikace zeptá na název a umístění výsledného souboru po konverzi. Po potvrzení názvu a umístění proběhne samotná konverze. O průběhu převodu je odečítač informován tzv. indikátorem průběhu.

Výsledný dokument se poté otevře, pokud je taková možnost pro převod nastavena. V případě převodu do formátu PDF se dokumentu otevře v interním editoru PDF.

5.4 Nová úloha - Skenovat

Jednou z nejčastěji používaných funkcí aplikace FineReader je úloha Skenovat s následným automatickým převodem naskenovaného obrazu do přístupné digitální podoby vhodné pro odečítače obrazovky.

Ve skupině Skenovat najdete tyto funkce:

- Skenovat: Editor OCR
- Skenovat: Dokument PDF
- Skenovat do: Microsoft Word
- Skenovat do: Microsoft Excel
- Skenovat do: Obrazové soubory
- Skenovat do: Ostatní formáty

Nejčastěji pravděpodobně použijete některou z prvních tří funkcí. Výběrem funkce ze skupiny Skenovat začne aplikace komunikovat s připojeným skenerem.

Pokud ještě není výchozí skener vybrán nebo není dostupný, zobrazí se dialogové okno pro výběr skeneru. Pokud se komunikaci se skenerem nepodaří navázat, bude úloha skenování pozastavena. Aplikace tuto skutečnost oznámí několika zprávami.

Po úspěšném navázání komunikace se skenerem zobrazí aplikace dialogové okno s možnostmi nastavení parametrů skenování. V dialogu prvky procházíte tabulátorem, prvky nastavujete k tomu určenými klávesami.

Důležité: Obecně pro skenování doporučujeme nastavit tyto parametry: rozlišení skeneru aspoň 300 dpi, skenovat v šedé barvě, jazyk podle automatického výběru, kvalita obrázku raději vyšší až nejvyšší.

Poznámka: Alternativně nastavíte parametry skenování také v editoru OCR v menu grafického okna („Zobrazit vlastnosti pole“ na kartě Vlastnosti obrazu) nebo v horním pruhu příkazů. Ostatní nastavení jsou dostupná v dialogu Možnosti.

5.4.1 Dialog Skenovat do editoru OCR

Nastavení v levém panelu okna

- Automaticky zpracovat obrázky stránek po jejich přidání >zaškrtnuto
- Přepínač Rozpoznat obrázky stránek >zapnutý
- Volba Jazyky OCR
- Volba Možnosti...

Skenování spustíte tlačítkem Skenovat, úlohu zrušíte tlačítkem Zrušit.

5.4.2 Dialog Skenovat do souboru PDF

Nastavení v levém panelu okna

- Kvalita obrázku >seznam Nejvyšší kvalita
- Vytvořit PDF/A >odškrtnuto; avšak zaškrtněte, jestliže pořizujete PDF pro někoho jiného (vysvětlení najdete v následující kapitole „Formát PDF“)
- Použít kompresi MCR >zaškrtnuto
- Rozpoznat text na obrázcích >zaškrtnuto
- Volba Jazyky OCR
- Volba Možnosti...

Skenování spustíte tlačítkem Skenovat do souboru PDF, úlohu zrušíte tlačítkem Zrušit.

5.4.3 Skenovat do MS Wordu

Nastavení v levém panelu okna

- Zachovat formátování >seznam Upravitelná kopie
- Volba Jazyky OCR
- Zachovat obrázky >zaškrtnuto
- Zachovat záhlaví, zápatí a čísla stránek >zaškrtnuto
- Volba Možnosti...

Skenování spustíte tlačítkem Skenovat do aplikace Word, úlohu zrušíte tlačítkem Zrušit.

Příkazem Skenovat naskenujete předlohu ve skeneru jako obrázek (bez rozpoznání textu). Po skenování se ukáže dialog Skenování obrázků se dvěma tlačítky „+Skenovat znovu“ a „Dokončit skenování“. Po „Dokončit skenování“ se automaticky spustí OCR a rozpozná v obrázku text. Indikátor průběhu informuje o postupu rozpoznávání OCR. Po dokončení se vypíše hláška „bylo rozpoznáno např. 2 z 2 stránek“ s tlačítkem Zavřít.

Při skenování do Wordu se po rozpoznání textu nabídne dialog pro uložení rozpoznání souboru ve formátu docx. Po uložení (zadáte název) se uložený soubor otevře v editoru Word, kde s ním můžete dále pracovat. Když okno Wordu minimalizujete nebo zavřete, ocitnete se v okně editoru OCR, které je popsáno v předchozí podkapitole Nová úloha – Otevřít.

5.5 Nová úloha – Porovnat

Úloha Porovnat umožňuje porovnat dvě verze dokumentu, i když jsou v různých formátech. Tato funkce není dostupná v některých verzích FineReader 15 či 14. Funkce je bohužel nepřístupná pro očečítače obrazovky, k jejímu použití je nutné ovládání myši a vizuální kontrola.

Otevřete úlohu Porovnat dokumenty, otevřete jednu z verzí dokumentu, kterou chcete porovnávat, v levém podokně a druhou verzi dokumentu v pravém podokně. Odklepněte tlačítko Porovnat a dokumenty porovnejte.

Nalezené rozdíly mezi porovnávanými verzemi budou v podoknech každé verze zvýrazněné a uvedené v pravém podokně, takže získáte jasnou představu o změnách, které byly v dokumentu provedeny. V obou verzích je tedy snadno vidět, jaký text byl přidán, odstraněn nebo upraven. Obě stránky se posouvají najednou a stejné části jsou vždy zobrazeny vedle sebe. Rozdíly lze odstraňovat ze seznamu nebo kopírovat do schránky.

5.6 Systémová integrace FineReaderu

Aplikace se při instalaci integruje se systémem, když během instalace přidává svoje rychlé úlohy do kontextového menu pro soubory zobrazené ve správcích souborů. Při instalaci jsou také vytvořeny ikony programu na pracovní ploše a panelu rychlého spuštění.

V kontextové nabídce pod položkou „Aplikace ABBYY FineReader“ či pod „Převést pomocí aplikace ABBYY FineReader“ jsou k dispozici funkce z dialogového okna Nová úloha. Nabízejí se zde funkce, které jsou použitelné pro daný zvolený typ souboru, například funkce pro převod do jiných formátů, případně jejich otevření přímo v aplikaci FineReader resp. otevření souboru v editoru OCR.

Dostupné možnosti zjistíte, když na nějakém souboru typu text, obrázek, prezentace, tabulka apod. vyvoláte kontextové menu. Vyberete možnost „Převést pomocí aplikace ABBYY FineReader“, načte se rozbalí možné formáty, do kterých lze soubor převést. Po volbě formátu už nemusíte program FineReader spouštět - k převodu dojde na pozadí a

po dokončení převodu se objeví obvyklý ukládací dialog, kde zadáte název a umístění výsledného souboru.

Poznámka: Volba z kontextového menu „Převést pomocí aplikace ABBYY FineReader“ bude patrně snazší a rychlejší, než stejná volba z prostředí aplikace FineReader.

6 Formát PDF

S dokumenty ve formátu PDF se můžete setkat poměrně často: články, manuály, návody, letáky bývají publikovány právě v tomto formátu. Jinak si vlastní dokument ve formátu PDF vyrobíte snadno sami, každý dokument můžete vytisknout jako PDF – v textovém editoru stačí vybrat tiskárnu „Microsoft Print to PDF“ a spustit tisk.

PDF je souborový formát vyvinutý firmou Adobe pro ukládání dokumentů nezávisle na softwaru i hardwaru, pomocí kterého byly pořízeny. Soubor typu PDF může obsahovat text, tabulky i obrázky, přičemž tento formát zajišťuje, že se libovolný dokument na všech zařízeních zobrazí a vytiskne stejně. Pro tento formát existují volně dostupné prohlížeče pro mnoho platforem, formát PDF otevřete např. v prohlížeči Edge nebo v oficiálním prohlížeči Adobe Reader, který lze stáhnout z <https://get.adobe.com/cz/reader/>.

Je dobré si uvědomit rozdíl oproti formátu HTML, který používají internetové stránky. Formát HTML si klade za cíl popsat obsah WWW stránky tak, že konkrétní způsob zobrazení stránky je na libovůli prohlížeče, potažmo uživatele. To na jedné straně umožňuje zobrazovat dokumenty tak, aby vyhovovaly čtenáři, na straně druhé znemožňuje zaručit stejný vzhled dokumentu u všech uživatelů. Oproti tomu cílem formátu PDF je zajistit co nejpřesněji stejné zobrazení na libovolném zařízení.

Formát PDF je proto před HTML upřednostňován zvláště v případě graficky bohatého nebo graficky přesného návrhu, u kterého autor vyžaduje dodržení původního vzhledu, např. u reklamních materiálů, prospektů, manuálů apod.

Typickým příkladem rozdílů mezi PDF a HTML je chování při zvětšování dokumentu, např. pro slabozrakého čtenáře:

- Při zvětšení dokumentu PDF se pouze zvětšuje měřítko, vzhled dokumentu zůstává zachován; při velkém zvětšení se tak na obrazovku nevejde celý řádek textu a je třeba text při čtení řádku vodorovně posouvat.
- Pro zvětšení HTML dokumentu stačí zvětšit velikost použitého písma, text se přeformátuje tak, aby se stále vešel do šířky okna prohlížeče, při tom se ovšem vzhled dokumentu změní, řádkové přechody jsou mezi jinými slovy, zvýší se počet řádků textu apod.

Formát PDF byl vyvinut pro výměnu dokumentů tak, aby zachoval a ochránil obsah a úpravu dokumentu (včetně jeho grafického uspořádání) a tak, aby vypadal vždycky

stejně bez ohledu na počítač či platformu. To je na druhé straně důvod, proč může být někdy obtížné PDF editovat nebo z něj vytahovat informace pro další použití.

6.1 Typy formátů PDF

Existují různé typy formátů PDF, se kterými je třeba také různě pracovat při vyhledávání nebo získávání textových informací. V zásadě způsob vzniku určuje, zda text PDF je přístupný či nedostupný. Podle způsobu vzniku lze rozlišit dva základní typy PDF:

- Přístupné či prohledávatelné formáty PDF, které jsou digitálně pořízené a vytvořené softwarem, např. ve Wordu příkazem Uložit jako (pro ukládání vyberete formát PDF) nebo funkcí Tisk na virtuální tiskárnu PDF (z menu Soubor > Vytisknout a zvolit tiskárnu „Microsoft Print to PDF“). Obecně takto vytvořené dokumenty PDF obsahují textovou i grafickou vrstvu, v textové vrstvě je veškerý text napsaný v editoru, v grafické vrstvě jsou pak obrázky, které jste do editoru vložili.
- Obrázkové (grafické) formáty PDF, které jsou naskenované nebo vyfotografované kamerou a poté uložené v některém grafickém formátu, např. jpg, tiff apod. Obsah takového grafického souboru je pak po vytištění virtuální tiskárnou PDF „uzavřen“ v PDF jako fotografie a případný text v obrázku je pro vyhledávání a odečítače nepřístupný.

Text v textové vrstvě PDF lze kopírovat a prohledávat (vyhledat v něm výskyt určitého slova) a je přístupný pro odečítače a pro hlasový výstup. Text uzavřený v obrázku přístupný není, lze jej číst pouze očima. Obecně lze říci, že nepřístupné nebo jen částečně přístupné jsou dokumenty, které obsahují významný text v grafické vrstvě (typicky naskenovaný obrázek s textem) nebo uzamčená PDF (chráněná heslem či kódovaná).

Dále ještě existují strukturovaná PDF (někdy také nazývaná tagovaná), která pomocí interních značek popisují strukturu dokumentu a pořadí různých prvků dokumentu (např. obrázky, bloky textu, sloupce, nadpisy). Označování usnadňuje extrakci textu nebo grafiky ze souborů PDF a pomáhá odečítacím programům i čtenářům vnímat obsah dokumentu ve správném pořadí, tedy např. správné pořadí čtení textu uspořádaného do sloupců.

Poznámka: Také se můžete setkat s formátem PDF/A, což je formát PDF pro archivaci elektronických dokumentů. Jeho norma zajišťuje, že soubory uložené v tomto formátu bude možné otevřít i v budoucnu a bude tedy kompatibilní s platformami, které se teprve objeví. Dokument PDF/A nesmí být chráněn heslem ani nesmí být zašifrován nebo zakódován, takže je v principu přístupnější.

Výhodou tisku PDF na virtuální tiskárnu je snadná dostupnost nástroje a možnost vytvořit PDF z jakékoliv aplikace, z níž lze tisknout. Nevýhodou pak je výsledné PDF na úrovni prostého textu, tzn. nestrukturované PDF. Proto při vytváření vlastních dokumentů PDF se spíše vyhněte tisku na virtuální tiskárnu PDF. Formát PDF vyrobený tiskem na

virtuální tiskárně příkazem „Print to PDF“ je přístupný odečítačům a odečítač jej dokáže přečíst, ale celková struktura originálu se může vytratit. Proto vždy, když je to možné, pro tvorbu PDF raději použijte příkaz „Uložit jako“ nebo „Exportovat“, které vytvoří strukturovaný dokument PDF.

S přístupnými PDF velké problémy nebývají, primárně je třeba řešit PDF nepřístupná nebo částečně přístupná. Taková grafická PDF bez textové vrstvy nebo s významným textem v obrázku nelze prohledávat a jejich text uzavřený v obrázku nelze upravovat.

V těchto případech je řešením technologie OCR: Po aplikaci OCR lze do grafického PDF přidat textovou vrstvu, která pak obsahuje text rozpoznáný v obrázku uzavřeném v PDF. Uvědomte si však, že převod OCR nemusí být vždy dokonalý a že tedy dokument po OCR nemusí plně odpovídat originálu.

Zda je daný dokument PDF přístupný či nikoli, poznáte velmi snadno: V přístupném PDF lze po otevření v Adobe Reader či Edge zvýraznit (vybrat) kurzorem blok textu, který lze pak vložit např. do schránky systému nebo do jiného textového editoru. Nebo lze zde také spustit hlasové čtení, které přečte textovou vrstvu PDF.

6.2 Asistivní čtečka PDF v prohlížeči Edge

V menu prohlížeče Edge najdete příkaz Číst nahlas (zkratka Ctrl+Shift+U), kterým spustíte čtení textové vrstvy otevřeného PDF hlasovým výstupem. Uprostřed horní lišty pod adresním řádkem je tlačítko Přehrát/Stop, kterým čtení ovládáte (nebo také již zmíněnou zkratkou Ctrl+Shift+U). Na liště vpravo pod tlačítkem Možnosti pro hlas si můžete nastavit rychlost čtení, případně vyměnit hlas za jiný. Asistivní čtečku zavřete křížkem „X“ na liště vpravo.

Pozor však, můžete se setkat s dokumentem PDF částečně přístupným, který obsahuje grafickou i textovou vrstvu a čtečka nebo příkaz Uložit jako text přečte resp. uloží pouze text v textové vrstvě. Případný text skrytý v obrázku (v grafické vrstvě) zůstane neviditelný a zpřístupní jej pouze proces OCR, který do textové vrstvy přesune veškerý rozpoznáný text.

6.3 Jak tedy přečíst nepřístupné PDF? Úkol pro OCR !

Jsou to stejné možnosti, které jsme probírali v kapitole „Jaké jsou prostředky pro OCR?“. Tedy v zásadě webová služba nebo specializovaný program OCR.

Webová služba: Nabídek služby, která převede dokument do přístupného formátu PDF pomocí technologie OCR existuje velmi mnoho, jsou bezplatné i placené. Stačí zadat příslušný požadavek do internetového vyhledávače.

Program OCR: Dále uvedeme dva způsoby - převod zabudovaný v odečítači Fusion/Jaws a program FineReader.

6.3.1 Fusion /JAWS: Převod OCR v dokumentu PDF

Otevřete soubor PDF a zadejte příkaz pro čtení formátu PDF: INSERT + MEZERNÍK, pak O, pak D. Fusion/JAWS řekne "OCR spuštěno" a otevře nové okno "Výsledky rozpoznávání textu", které už čtete běžným způsobem.

6.3.2 FineReader – Nová úloha: Otevřít dokument PDF

Dokument PDF otevřete v editoru PDF z okna Nová úloha, které je popsáno v kapitole „Rozpoznávací program ABBYY FineReader PDF“.

V hlavním okně Nová úloha zvolte Otevřít > Dokument PDF (nebo také Nová úloha > Soubor > Nový dokument PDF či Otevřít dokument PDF). Otevře se standardní dialogové okno pro výběr existujícího dokumentu z úložného prostoru počítače nebo z jiného dostupného úložiště. Zvolený dokument se pak zobrazí v editoru PDF, v podokně Dokument, kde lze text pouze číst. Chcete-li text dokumentu upravovat v interním editoru, musíte z okna Dokument přepnout do okna Upravit obsah.

Editor PDF obsahuje několik funkcí pro práci s dokumentem jako celkem nebo s jeho obsahem. Dostupné funkce najdete v kontextové nabídce nebo v hlavním menu aplikace. Hlavní menu otevřete pomocí klávesových zkratk (ALT a ŠIPKA NAHORU/DOLŮ, případně ALT+písmeno pro jednotlivé skupiny funkcí). Každá skupina funkcí v hlavním menu má svou vlastní klávesovou zkratku, kterou odečítáč sdělí při přechodu mezi skupinami. Mezi skupinami a v jejich nabídce se pohybuje pomocí šipek.

Režim celé obrazovky

Editor PDF má k dispozici režim celé obrazovky, v němž je dokument zobrazen na celé obrazovce a nejsou zobrazeny žádné panely ani panely nástrojů. Chcete-li přejít do režimu celé obrazovky, proveďte jednu z následujících akcí:

- Odklepněte tlačítko na panelu nástrojů ve spodní části obrazovky
- V hlavní nabídce vyberte Zobrazit > Celá obrazovka
- Stiskněte klávesu F11

Funkce dostupné v tomto režimu najdete v kontextovém menu zobrazeného dokumentu. Jinak, jestliže chcete zobrazit nástroje v panelu nástrojů, přesuňte ukazatel myši na spodní okraj obrazovky.

Režim celé obrazovky ukončíte klávesou F11 nebo ESC.

Rozpoznávání na pozadí

Editor PDF umožňuje hledání textu a obrázků v dokumentech PDF bez textové vrstvy a kopírování textu a obrázků z nich – jde například o naskenované dokumenty a dokumenty vytvořené z obrazových souborů. To je možné díky tomu, že proces OCR

běží na pozadí. Ve výchozím nastavení je povoleno rozpoznávání na pozadí a spouští se automaticky, když se otevře dokument PDF.

Proces rozpoznávání na pozadí nemění obsah originálního souboru PDF. Místo toho přidá dočasnou textovou vrstvu, která nebude při otevření dokumentu v jiných aplikacích k dispozici. Chcete-li, aby dokument bylo možné prohledávat i v jiných aplikacích, budete muset textovou vrstvu vytvořenou při procesu rozpoznávání na pozadí uložit. Chcete-li tak učinit, klepněte na tlačítko na panelu nástrojů na kartě Dokument a z rozevřacího seznamu vyberte „Rozpoznat dokument...“.

Důležité: Pokud funkce hledání a kopírování nefungují správně, prověřte, zda jsou pro dokument správně zadány jazyky OCR.

7 Bonusy

7.1 Malý test přístupnosti a nepřístupnosti PDF

Na stránce <https://spektra.eu/vyuka/zoomtext-tutorialy/> v místě, kde jste si stáhli tuto lekci tutoriálu, najdete tři příklady souborů ve formátu PDF:

- PDF-přístupný a nepřístupný_netagovaný, pořízený tiskem na virtuální tiskárně PDF z Wordu (je netagovaný),
- PDF-přístupný a nepřístupný_tagovaný, pořízený příkazem "Uložit jako" ve formátu PDF z Wordu a jako tagovaný,
- PDF-přístupný a nepřístupný_heslo, pořízený příkazem "Uložit jako" ve formátu PDF z Wordu a uzamčený heslem a kódovaný. Heslo pro otevření této ukázky je QWERTZ.

Obsah všech tří souborů je stejný, každá ukázka PDF obsahuje obě vrstvy, významný text v obrázku a text v textové vrstvě, který je navíc uspořádaný do tří sloupců.

Když otevřete první soubor (netagovaný) v prohlížeči Edge, snadno se můžete přesvědčit, že text v obrázku přístupný není, zatímco text v textové vrstvě můžete kopírovat, prohledávat (zkratka Ctrl+F) nebo si jej nechat přečíst zabudovanou čtečkou (zkratka Ctrl+Shift+U).

Když v prohlížeči Edge otevřete třetí soubor (uzamčený heslem), zjistíte, že se přístupnost znatelně zhoršila - zabudovaná čtečka se na textu v textové vrstvě zasekává či úplně přerušuje.

Formát PDF si také můžete otevřít ve Wordu, kde si u jednotlivých ukázek ověříte přístupnost textu v textové vrstvě pro odečítače. Chování bude podobné jako v prohlížeči Edge. Znatelný rozdíl zjistíte při otevření souboru PDF ve Wordu:

- tagovaný formát PDF se převede přesněji a lépe odpovídá formátování originálu,

- zatímco netagovaný formát PDF se převede hůře a jeho formát je ve srovnání s originálem rozhozenější.

7.2 Jak lze získat strukturované PDF

7.2.1 Strukturované (tagované) PDF při ukládání či exportu z Wordu:

Z menu Soubor > Uložit jako > Počítač > a tlačítko Procházet otevře standardní dialog pro uložení souboru, v dialogu vyberete formát PDF. Pod dialogem je tlačítko Možnosti, po jeho otevření by měla být zaškrtnutá položka „Značky struktury dokumentů pro usnadnění“ a nastavení potvrdíte OK.

Z menu Soubor > Exportovat > Vytvořit dokument PDF > a tlačítko Vytvořit soubor PDF otevře standardní dialog pro uložení souboru ve formátu PDF. Pod dialogem je tlačítko Možnosti, po jeho otevření by měla být zaškrtnutá položka „Značky struktury dokumentů pro usnadnění“ a nastavení potvrdíte OK.

7.2.2 Archivní formát PDF/A podporuje také Microsoft Office od verze 2010:

Uložit jako > tlačítko Procházet a otevře se standardní dialog pro uložení souboru, zvolte Uložit jako a příponu PDF. Pod dialogem je tlačítko Možnosti: po jeho odklepnutí v sekci Možnosti PDF zaškrtněte položku "Kompatibilní s normou ISO 19005-1 (PDF/A)" a potvrdíte OK. Uložené PDF nyní bude odpovídat archivnímu standardu PDF/A.

7.3 Jak převést dokument PDF do formátu Wordu?

Jednoduchý převod z PDF do Wordu, který by fungoval univerzálně a bez potíží, neexistuje. Chcete-li převést PDF do formátu DOC/DOCX s co možná nejlépe zachovaným vzhledem stránky, uděláte nejlépe, použijete-li OCR program, který si s převodem poradí.

Nejlépe asi funguje tento převod v ABBYY FineReaderu. Adobe Reader sice nabízí export z PDF do různých formátů, včetně DOC, kdo však někdy tuto funkci zkoušel, ví, jak nedobře to funguje. Formát PDF lze také otevřít přímo ve Wordu z menu Soubor > Otevřít, převede se však pouze textová vrstva PDF a formátování nemusí úplně odpovídat originálu.

7.4 Jak nouzově na nepřístupná PDF

I s nepřístupnými dokumenty PDF si lze poradit: V některých případech lze takový dokument PDF přímo rozpoznat OCR programem, krajní řešení pak obvykle spočívá ve vytištění dokumentu, jeho naskenování a rozpoznání OCR programem. Takto prolomíte i uzamčená či zakódovaná PDF, která v digitální podobě mohou vzdorovat i specializovaným programům.

7.5 Rozpoznání obrazovky: ScreenShot Reader

Screenshot Reader, tedy čtečka sejmuté obrazovky, je nástrojem FineReaderu, který najdete v jeho hlavním menu ve skupině Nástroje. Nástroj Screenshot Reader sejme obsah obrazovky a spustí OCR na pozadí (převéde grafiku na text) a rozpoznaný text zkopíruje do schránky.

Odečítač Fusion/Jaws má také zabudovanou funkci „Provést OCR na snímku obrazovky“, stačí zadat příkaz INSERT + MEZERNÍK, pak O, pak S.

8 Některé klávesové zkratky z této lekce

Kamera v ZoomTextu

Příkaz	Klávesová zkratka
Spustit kameru v ZoomTextu	CAPSLOCK + CTRL + C
Zvětšit / zmenšit	Šipka NAHORU / DOLŮ
Další funkce kamery	Myši z aktivní nástrojové lišty Kamery
Ukončit kameru v ZoomTextu	Zavřít nástrojovou lištu Kamery

Hlasová čtečka v prohlížeči Edge

Příkaz	Klávesová zkratka
Čtečka v prohlížeči Edge: Zap/Vyp	CTRL + SHIFT + U

OCR v odečítači Fusion/Jaws

Příkaz	Klávesová zkratka
Fusion/Jaws: OCR do Wordu	INSERT + MEZERNÍK, pak O, pak R
Fusion/Jaws: Spustit čtení PDF	INSERT + MEZERNÍK, pak O, pak D
Fusion/Jaws: OCR na snímku obrazovky	INSERT + MEZERNÍK, pak O, pak S

Nová úloha (FineReader)

Příkaz	Klávesová zkratka
Nový dokument PDF	CTRL + N
Otevřít dokument PDF	CTRL + O
Otevřít v editoru OCR	CTRL + R
Skenovat do editoru OCR	CTRL + S

Editor OCR (FineReader)

Příkaz	Klávesová zkratka
Okno Nová úloha	CTRL + N
Otevřít obrázek	CTRL + O
Skenovat stránky	CTRL + K
Zobrazit podokno Navigace (Stránky)	F5
Zobrazit podokno Lupa	CTRL + F5
Zobrazit podokno Obraz	F6
Zobrazit jen okno Text	F8
Zvětšit / Zmenšit	CTRL + Plus / CTRL + Minus
Rozpoznat stránku / všechny stránky	CTRL + R / CTRL + SHIFT + R

Editor PDF (FineReader)

Příkaz	Klávesová zkratka
Okno Nová úloha	CTRL + N
Otevřít PDF	CTRL + O
Zvětšit / Zmenšit	CTRL + Plus / CTRL + Minus
Celá obrazovka	F11

Nenechte se deprimovat seznamem klávesových zkratk a nezoufejte: Naučte se jen ty klávesové zkratky, které budete častěji používat - zrychlí vám práci. Nemusíte si dokonce pamatovat žádné zkratky, vždy lze odpovídající akci provést z menu nebo příkazem v pásu nástrojů, jenom je to pak trochu delší, na více stisků.

9 Jak dále

Se znalostmi, které jste získali v této lekci, se už dokážete poměrně dobře orientovat v technologii skenování a OCR a v práci s programem FineReader. Nezapomínejte ani na možnost doplnit si podle potřeby tento kurz další náповědou zejména k programu FineReader, který toho umí více, než jsme si v této lekci ukázali.

Upozorňujeme také na stránku <https://spektra.eu/vyuka>, kde najdete další výukové materiály a další informace.

A nebojte se počítače: je to velmi snadné ...

9.1 Licence

Tento kurz je určen pro výuku zrakově postižených uživatelů speciálního zvětšovacího a odečítacího programu ZoomText nebo Fusion a autoři kurzu nabízejí pro tento účel celý text nebo jeho části k bezplatnému použití a volnému šíření. Zároveň však autoři prosí, aby přitom vždy byl uváděn odkaz na tento původní zdroj, z něhož použití textu kurzu nebo jeho částí pochází. Pro jiný účel než výše uvedený, nelze text kurzu nebo jeho části bez souhlasu autorů použít, v žádném případě jej nelze šířit za úplatu.

9.2 Poděkování a kontakt

Autoři velmi děkují za všechny připomínky a návrhy na doplnění kurzu, které chápou jako společnou snahu pomáhající ke zlepšení výuky uživatelů uvedených produktů. Případné připomínky či návrhy k obsahu kurzu zasílejte prosím redaktorovi kurzu na adresu: verner.bretislav@spektra.eu.

Zpracování kurzu bylo částečně sponzorováno družstvem Spektra, v.d.n., které je zároveň autorizovaným prodejcem předmětných produktů (www.spektra.eu).

