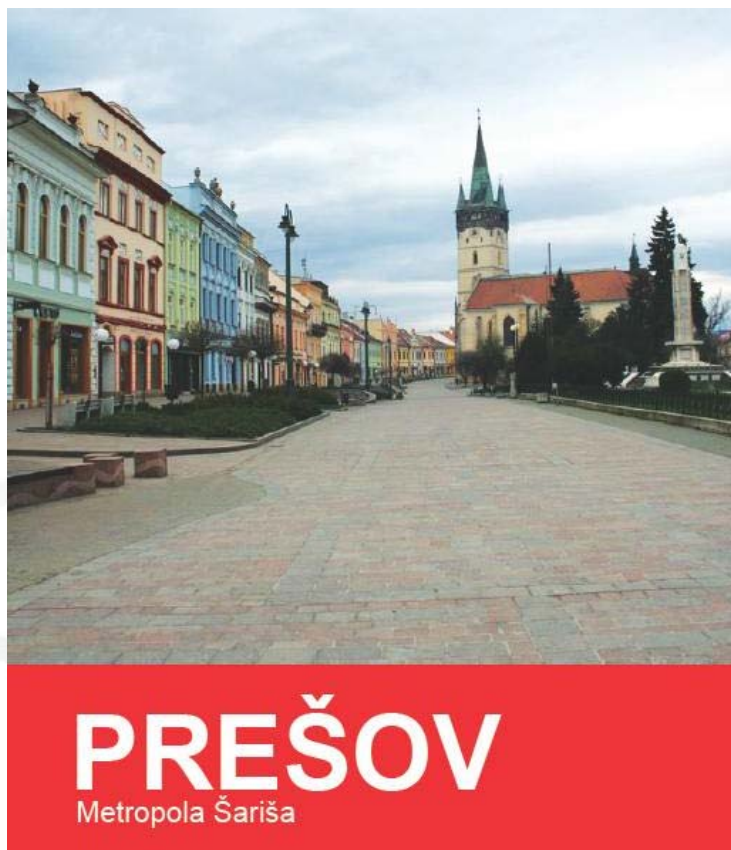


Konverzia dokumentov do elektronickej podoby



Dom kultúry Devín
Bratislava - Devín

Konverzia dokumentov do elektronickej podoby

Základné požiadavky pre text v elektronickej dokumente

(usmernenie MF SR pre konverziu dokumentov)

- ❖ čitateľný,
- ❖ označiteľný a kopírovateľný (pokiaľ nie je kopírovanie z odôvodnených dôvodov nepovolené),
- ❖ manuálne vyhľadávateľný (používateľom),
- ❖ spracovateľný inými systémami (napr. kvôli indexácii, vytváraniu vyhľadávacích kategórií, samotnému vyhľadávaniu atď.)

Nedodržanie týchto požiadaviek zabraňuje ďalšiemu spracovávaniu, vyhľadávaniu v texte alebo čítaniu textu zrakovo postihnutými, čo je v rozpore s princípmi informatizácie spoločnosti.

Proces konverzie - skenovanie

Požiadavka - uložené skenované textové dokumenty nesmú mať v zmysle štandardov podobu obrázkov - musia byť „textovo rozoznatel'né“.

Spôsob - použitie OCR (Optical Character Recognition) s rozoznáváním slovenských znakov

Výnimky - obrázky, logá, pečiatky, symboly, ozdobný text, text písaný rukou
- nemusia vyhovovať požiadavke „textovej rozoznatel'nosti“
- môžu mať naďalej podobu obrázkov

Odporúčané minimálne technické parametre pre skenovanie

- rozlíšenie min. 300 DPI
- minimálna veľkosť obrázka 600x800 bodov pri skenovaní formátu A4

Nepovinné odporúčania pre zverejňovanie dokumentov

Použiteľné pre všetky typy dokumentov, nie len pre skenované

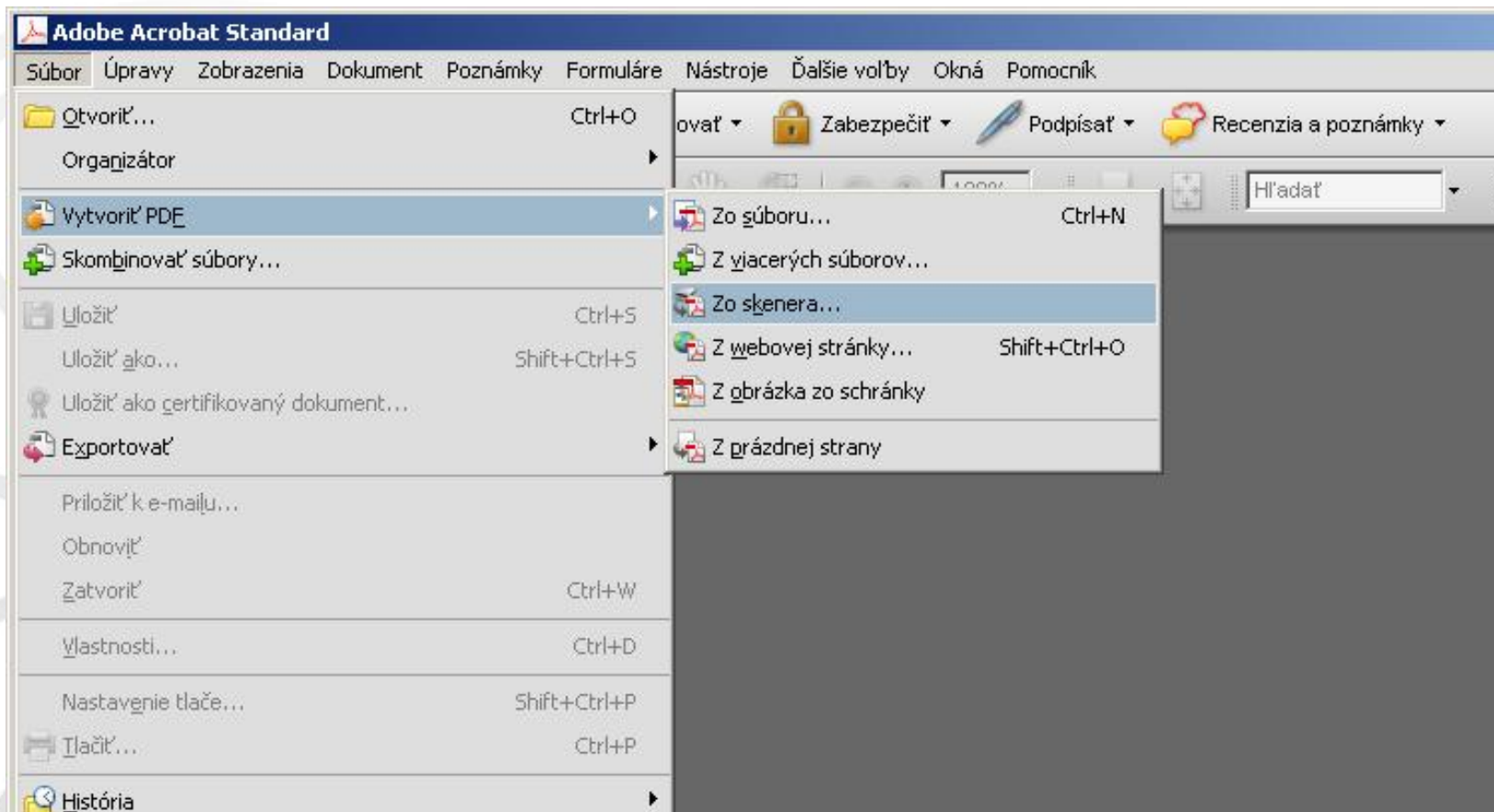
- ❖ Pri zverejňovaní súborov používať štandardné prípony súborov, aby nevznikali problémy pri identifikovaní ich formátu a pri ich otváraní.
- ❖ Užitočné je pri zverejňovanom súbore uvádzať aj dátovú veľkosť súboru a jeho formát. V prípade použitia ikoniek pre identifikáciu typu súboru je vhodné na webových stránkach používať také, ktoré nesymbolizujú konkrétnu komerčnú aplikáciu.
- ❖ Názvy zverejňovaných súborov čo najviac skrátiť - optimálne je obmedziť ich dĺžku do 64 znakov.
- ❖ V názvoch zverejňovaných súborov je z hľadiska prístupnosti z rôznych platforiem optimálne používať len písmená bez diakritiky, ďalej čísla a znaky podčiariť, mínus, plus (_ - +).
- ❖ Používanie diakritických znamienok, medzier a najmä špeciálnych znakov (/ \ : * ? " < > |) v názvoch súborov môže sťažiť alebo úplne znemožniť prístup k obsahu súboru.
- ❖ Pre pomenovávanie zverejňovaných súborov je vhodné prijať vnútorný predpis

Zdroj: Podporná stránka Výnosu o štandardoch pre ISVS. Stránku spravuje [Slovenská infromatická spoločnosť](#). Aktualizácia: Október 2010

Praktický príklad

Skenovanie dokumentov z aplikácie Adobe Acrobat 8.0

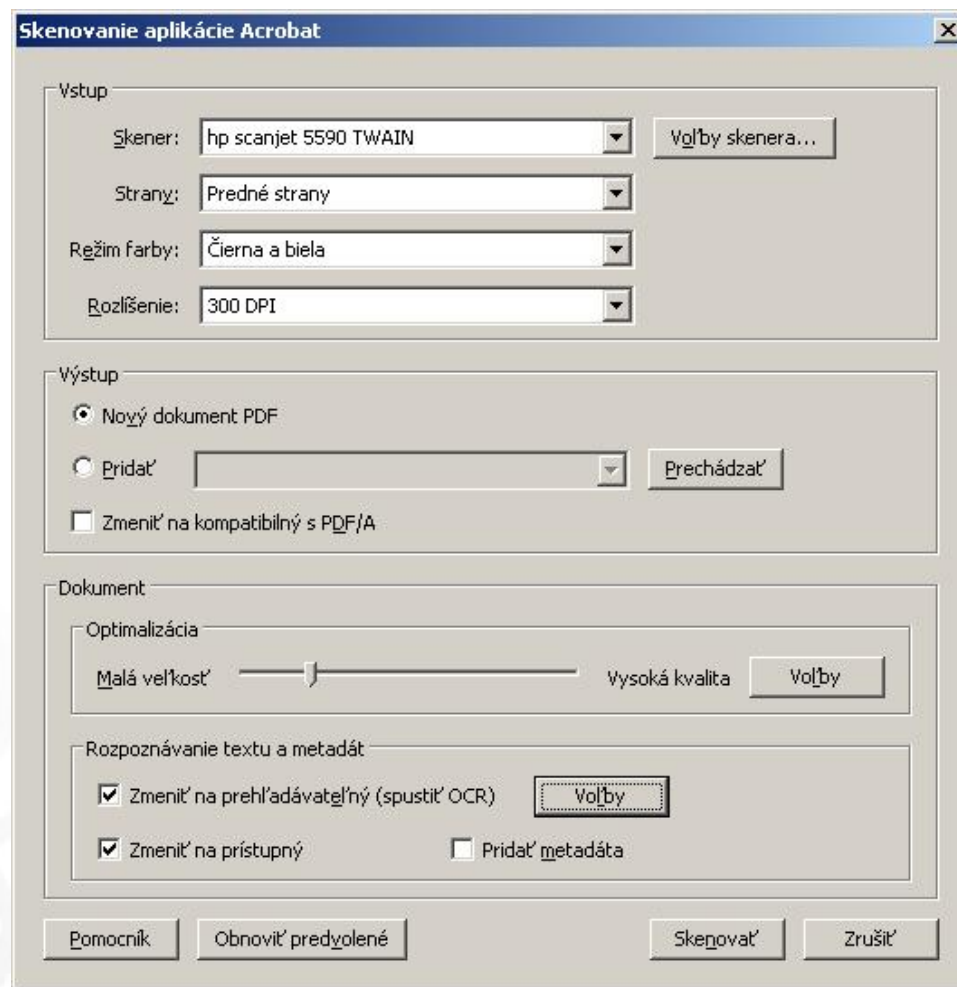
- Z menu vybrať *Súbor -> Vytvoriť PDF -> Zo skenera*



Praktický príklad

Skenovanie dokumentov z aplikácie Adobe Acrobat 8.0

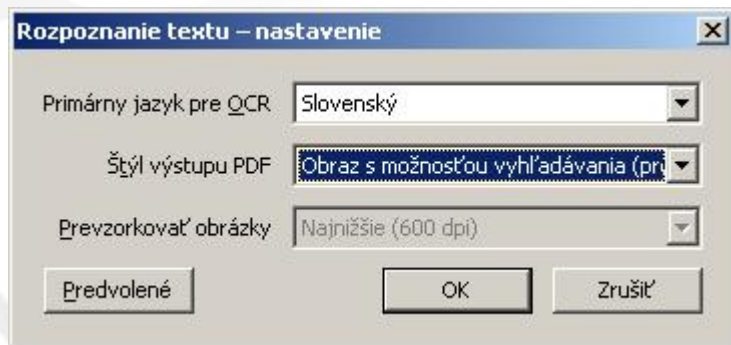
- Nastavenie parametrov pre skenovanie



Praktický príklad

Skenovanie dokumentov z aplikácie Adobe Acrobat 8.0

- Nastavenie pre rozpoznávanie textu: toto okno sa zobrazí pri stlačení tlačidla Voľby v časti Rozpoznávanie textu a metadát



- Obraz s možnosťou vyhľadávania (presný) Zabezpečí, že text možno prehľadávať a vyberať. Táto voľba zachová pôvodný obrázok a umiestni cez obrázok neviditeľnú vrstvu textu. Odporúča sa pre prípady, keď je potrebná maximálna vernosť pôvodného obrázka
- Najlepší text pre prevod vznikne čiernobielym skenovaním s rozlíšením 300 dpi

Priebeh konverzie

aní súborov je vhodné používať štandardné prípony súborov, aby pri identifikovaní ich formátu a pri ich otvárani.

pri zverejňovanom súbore uvádzať aj dátov
použitia ikoniek pre identifikáciu typu súboru je vhodné na webových



- Časová náročnosť konverzie 7 strán cca 3 min.
- Časovo náročnejšia je príprava dokumentu pred skenovaním (vybielenie nezverejniteľných textov), po konverzii - pomenovanie dokumentu a jeho publikovanie na internete.



Ďakujem za pozornosť

Kontakt:

mariana.hurna@presov.sk
miroslav.novak@presov.sk