# Návod na tvorbu obrázků a na ocerizaci ze skenů (1. fáze zpracování písniček zadaného alba)

## 2. Tvorba prohledávatelného PDF a RTF pro P a S verzi

1. Ořezaný text písně (viz tvorba obrázku ve formátu .tif, viz Návod na výřez písně) otevřeš v **ABBYY FineReaderu** (ke stažení např. : <https://www.stahuj.cz/grafika_a_design/fonty/programy_pro_praci_s_fonty/abbyy-finereader/>)
	1. Je tam 2 týdenní zkušební verze, kterou lze v případě problémů s bodem b. využít alespoň provizorně
	2. pro plnou verzi je potřeba mít zřízený vzdálený přístup (VPN) k počítačům v pracovně doktorandů na Ústavu románských jazyků a literatur (ten by měli mít všichni studenti kurzu FJ0B801 v PS2020 už zařízený, pokud nefunguje, ozvat se doc. Polické obratem), po skončení karantény bude možné pracovat i přímo z těchto počítačů na MU. Počítače jsou označeny jako ROMAN26, ROMAN27, ROMAN28 a ROMAN29. Návod na připojení VPN je zde: <https://it.muni.cz/sluzby/terminalove-servery/navody/pripojeni-k-serveru-ucn-tserver>, po připojení k jednomu z počítačů je potřeba vybrat program ABBYY FineReader, který je licencován pro souběžné použití nejvíce 4 studentů, tudíž může být v daný moment vytížen, vždy je to ale jen na chvíli, takže pokud připojeni jste a nefunguje program, patrně stačí jen přijít o chvilku později a bude fungovat – bude doupřesněno!
2. Nejdřív si obrázek ve formátu .tif převeď do vyhledávatelného PDF
	1. SOUBOR 🡪 NOVÝ ÚKOL 🡪 PŘEVÉST DO PDF
	2. V tabulce v novém okně, který se objeví
		1. Nejvyšší kvalita
		2. Fulltextové vyhledávání 🡪 první možnost: vyhledávat v textu a v obrázcích
		3. Check: vytvořit PDF.A
		4. Check: povolit kompresi (může a nemusí být, pokud ano, zkontroluj výsledek, jestli je v dobré kvalitě)
		5. Jazyk: FR
	3. Převést do PDF
3. Ocerizace
	1. SOUBOR 🡪 OTEVŘÍT OBRAZ, vybrat všechny JPG, co chceš zpracovávat
	2. Zkontroluj, že nahoře máš jako jazyk nastavenou FR, dost ti to pak zjednoduší další kontrolu, protože to automaticky rozpoznává fr slova
	3. Jakmile proběhne automatické rozpoznání, zkontroluj spodní dokument, jestli program rozpoznal všechno
		1. Pokud ne, dooznač a dej nahoře „ROZPOZNAT“
		2. Pokud ani tak nerozeznal, dopiš do okna vpravo ručně
		3. V pravém okně uprav všechno, co program označil špatně, popř. dopiš, co program nevzal
		4. Zkontroluj vše pečlivě řádek po řádku
		5. Když máš hotovo, můžeš dokument uložit ve formátu RTF (soubor 🡪 uložit jako 🡪 dokument RTF) a pak už upravovat ve wordu, nebo to přes ctr + c projet přes poznámkový blok, např. když se potřebuješ zbavit např. černého pozadí, které ti to vygeneruje podle bookletu



* 1. Ať tak jako tak, RTF dokument musí mít:
		1. Arial, velikost 10 (dobré projet kontrolu nejdřív přes Times New Roman, jsou tam vidět rozdíly, např. I/l 🡪 l/I)
		2. Řádkování 1,5
		3. Znovu překontroluj hlavně apostrofy, musí odpovídat tomu, co je na fr. klávesnici pod 4 ( ‘ )
		4. Překontroluj i s poslechem hudby (klidně už dřív, pokud budeš „slepovat“ sloky)
		5. Uprav tak, aby se RTF verze shodovala s verzí v bookletu (.tif) a PDF (viz další obrázek, vlevo: TIF, vpravo: RTF), to znamená dodržovat řádky a přepisovat i s případnými chybami!



1. Ulož do ISu pod názvem kódu písně formáty RTF a PDF do složky **Data FR písně 🡪 AA texty OCR (dvouvrstvá PDF)**
	1. Až když máš všechno hotové 🡪 ve složce bude sudý počet dokumentů
	2. Ukládej pod kódem, jaký byl i na vstupním TIFu. Opět platí, že zaškolovací píseň musí nejprve schválit kontaktní osoba, až pak vkládej samostatně.
	3. Každé dva dokumenty (s totožným kódem, jen v jiném formátu .pdf nebo .rtf) si musí obsahově a v rámci možností i formálně (až na font, velikost písma. atd., které jsou standardizované pro naše OCR dokumenty) odpovídat!!
	4. IS si sám automaticky generuje další verze – texťáky, word dokumenty, dokumenty, PDF, o to ty se nemusíš starat a nemělo by Tě to mást, vkládáš jenom tyhle dva formáty – PDF + RTF.
2. Zapiš si dobu práce do výkazu.

Tipy:

* Průběžně zapisuj odpracované hodiny a stav rozpracovaných písní, určitě si je už pouštěj do sluchátek, ať se s nimi lépe seznámíš.
* Je vhodné mít vždy aspoň 2 alba zadané do zásoby dopředu.

Kontaktní osoba pro PS 2020: **Adélaïde Evreinoff** <evreinoff.adelaide@gmail.com> (UČO 464372)