

Projekt ISJ

Automatické zarovnanie titulkov

Martin Krippel

xkripp00

Verzia Python: python 2.7.6

Skúšané na:

- merlin.fit.vutbr.cz
- vlastný notebook: Windows 7 Professional, 64-bit

Použité knižnice: urllib2, re, sys, getopt, os

Postup programu:

Program na základe zadanej URL nájde odkaz z danej stránky na súbor s titulkami a následne stiahne české titulky. Potom zistí ID filmu. S pomocou ID nájde anglické titulky. Všetky stiahne. Vyberie si dvojicu titulkov s najmenším rozdielom v počte riadkov. Pomocou regulárnych výrazov spracuje titulky. Následne sa ich snaží priradovať.

Kvôli problém s priradovaním titulkov je výsledné priradenie dosť zlé. Nepodarilo sa mi prísť na systém, ako titulky správne vyberať. Problém bol aj v podcenení projektu (resp. náročnosti).

Nakoniec program vytvorí súbor `vysledok.txt`, v ktorom sú priradené titulky. Program zmaže všetky titulky ktoré stiahol, ale nepoužil. Po skončení zostanú len titulky, ktoré sa program rozhodol zarovnať a súbor so zarovnanými titulkami.