

1) ÚVOD

Implementační dokumentace popisuje řešení 2. projektu do předmětu IPP. Vybral jsem si zadání CSV(CSV2XML v Pythonu). Úkolem byla implementace programu pro konverzi formátu CSV do XML v jazyku Python . Podobu výstupního XML lze ovlivnit parametry při spuštění programu.

2) POPIS ŘEŠENÍ

a) Zpracování parametrů příkazové řádky

Funkce pro zpracování parametrů příkazové řádky, je volána ihned na úvod programu, kontroluje správnou kombinaci zadaných parametrů a ukládá je do vhodné struktury, během programu jsou tak ihned k dispozici. U parametrů, z kterých je potřeba získat nějaké informace jako vstupní soubor apod., je provedeno ořezání na potřebnou informaci. Pro výpis nápovědy s popisem všech parametrů se program spouští s parametrem „--help“.

b) Vstup a výstup programu

Vstup je implicitně nastaven na čtení ze standardního vstupu, při zadaném parametru se vstupním souborem, dochází k ověření existence souboru. Výstup automaticky probíhá na standardní výstup, při sepnutém parametru s jménem výstupního souboru dojde k výstupu do tohoto souboru.

c) Konverze CSV do XML

I.) Bez hlavičky XML

Před samotným zpracováním CSV vstupu se na základě bool hodnoty proměnné uchováující informaci o sepnutém parametru tiskne nebo netiskne hlavička XML.

II.) Element obalující XML

Podle parametrů příkazové řádky je výstupní soubor XML obalen párovým elementem, jehož jméno je specifikováno z parametrů příkazové řádky. Při této akci dojde také k nastavení odsazení na řádku do příslušné proměnné.

III.) Separator

Pro vstupní soubor CSV může být specifikován separátor sloupců. Implicitně je nastaven na znak čárka (.). Pro načítání CSV souboru jsem použil csv modul a ten umožňuje nastavení oddelovače. Specialitou je oddělovač „TAB“, při něm je nastaven oddělovač na tabulátor.

IV.) První řádek CSV je hlavičkou

Na prvním řádku se nachází hlavička CSV, tedy názvy sloupců. Dochází ke zpracování 1.řádku vstupu a uložení názvu sloupců do vhodné struktury. Načtené údaje je potřeba zvalidovat, dochází k nahrazení znaků jako čárka, tečka nebo mezera. Další řádky jsou již zpracovány běžnou konverzí do XML formátu s použitím názvů sloupců načtených z prvního řádku CSV vstupního souboru.

V.) Element obalující řádek

Tag obalující řádek vstupního CSV souboru je implicitně s názvem „row“. Při zadaném parametru může být název změněn na řetězec z parametru.

VI.) Atribut elementu řádku

S elementem pro obalení řádku může být vytisknut číselný atribut, který bude zvyšován při každém použití tohoto elementu.

d) Zotavení z chybného počtu sloupců

Pokud počet sloupců ve vstupním souboru CSV není stálý, dojde k chybě, smazání výstupu a skript končí s výpisem na chybový výstup. Se zadaným parametrem pro zotavení z chyb, však program doplní chybějící sloupce.

3) ZÁVĚR

Skript provádí vše podle zadání, byl testován poskytnutými vzorovými testy a také testy vlastními. Výsledky testů prokázali správnost řešení.