

Fulltextové vyhledávání

Mějme databázi článků s id, titulkem a textem.

Úkolem je napsat program, který zadané slovo vyhledá v databázi a vypíše id a titulek článku, ve kterém se slovo nachází.

Je nežádoucí procházet při každém dotazu celou databázi znova. Lepší cesta je pomocí tzv. indexování: program články projde jednou a zapamatuje si mapu (index, rejstřík) slov a jejich výskytů v článcích.

Úkoly v bodech: - navrhněte si datovou strukturu (tj. zvolte vhodný datový typ, jednoduchý či složený), která reprezentuje články - navrhněte datovou strukturu indexu - program by se měl dotázat uživatele na slova (**input**), která chce vyhledat a vrátit mu výsledky

Ukázka:

Search for: mugabe plastic

[MGB01] Mugabe sacks 'disloyal' Vice-President Mnangagwa
Mugabe is expected to be appointed Vice-President ...

[PLP04] Plastic problem being discussed
plastic in the world's oceans so now experts are t...

no search results for africa

Poznámky: - přiložen testovací soubor ("databáze") s několika články. - neřešte apostrofy - vypište pouze první výskyt slova a následujících N znaků