

Projekt bazy danych

Przetwarzanie dokumentów XML

Adam Jędrychowski

23 czerwca 2023

1 Sekcje:

1. Cel projektu
2. Python jako biblioteka
3. Uruchomienie
4. Opis funkcjonalności API
5. Testy

2 Cel projektu

Stworzenie API, które pozwoli na łatwe przetwarzanie dokumentu xml w bazie danych.

3 Python jako biblioteka

Biblioteka ma na celu implementować zbiór funkcji i klas, które służą do ściśle określonego celu. Jeśli chcemy, aby z danej biblioteki można również korzystać w innych językach programowania niż ten w którym jest napisana musimy skorzystać z bibliotek konkretnego języka. Jeśli chcemy skorzystać z pliku dll w pythonie to użyjemy biblioteki ctypes, która udostępnia API do uruchamiania funkcji zawartych w dll. Z kolei jeśli chcemy użyć pliku .py jako biblioteki w c++ to musimy skorzystać z API udostępnianego poprzez bibliotekę Python. Podsumowując, aby połączyć dowolne 2 języki programowania korzystamy z API udostępnianego przez dany język.

4 Uruchomienie

Do uruchomienia aplikacji potrzebny jest python 3. Wymagana jest również instalacja xmlAlchemy i pyqt6 (pip install ...). Jeśli chcemy uruchomić testy to trzeba doinstalować unittest.

5 Opis funkcjonalności API

Biblioteka tworzy w bazie danych 2 relacje. Pierwsza przechowuje id węzłów, nazwę, id rodzica oraz numer kolejności. Druga przechowuje atrybuty węzłów, czyli id węzła, klucz i wartość (pole tekstowe ma klucz `__text__`). Biblioteka umożliwia:

- zapis i odczyt xml w bazie
- wyczyszczenie wszystkich danych
- usunięcie danego dokumentu xml
- wypisanie dostępnych dokumentów
- pobranie id węzła po jego numeru lini
- aktualizacja wartości atrybutów węzłów
- wstawienie dokumentu xml jako dziecko danego węzła
- zamiana kolejności węzłów tego samego rodzica
- znalezienie wartości atrybutu w dokumencie

Opis konkretnych funkcji znajduje się w kodzie.

6 Testy

Każda funkcja, która jest endpointem jest testowana w pliku `xml_tests.py`. Przyjmowany xml może się różnić znakami `"`, `'`, następna linia, dlatego w wielu testach xml wczytany z bazy jest normalizowany.