

Python średnio zaawansowany

Dzień 10



Blok nr 3:

Przetwarzanie danych

AGENDA

- Powtórka wiadomości nt ORM
 - Powtórka wiadomości nt SQLAlchemy
 - Kilka trików, które warto znać
 - Omówienie rozwiązań zadania domowego
 - Wasze najciekawsze doświadczenia
-
- ... a to dlatego, że dane z bazy
będziemy wykorzystywali w kolejnym bloku kursu

Object Relational Mapping

ORM - definicja

Mapowanie obiektowo - relacyjne polega na **mapowaniu** danych w postaci **tabelarycznej** (relacji w bazie danych) na **obiekty** i odwrotnie.

To nowoczesne podejście do zagadnienia współpracy z bazą danych, wykorzystujące filozofię programowania obiektowego.

Zalety ORM

- Niezależność od silnika bazodanowego
- Większa produktywność programisty
- Nadal jest możliwość wstrzyknięcia kodu SQL jeżeli ten generowany przez ORM jest zbyt skomplikowany/wolny

Wady ORM

- Skomplikowane zapytania SQL będą wymagać zapisania ich wprost, bez wykorzystania ORM
- Mechanizm ORM może być powolny

SQLAlchemy

SQLAlchemy

- Instalacja:

```
pip install sqlalchemy
```

- 4 minimalne wymagania dla modelu (klasy mapowanej na tabelę):
 1. dziedziczy po `declarative_base`
 2. nazwa tabeli zapisana w `__tablename__`
 3. posiada przynajmniej jedną kolumnę
 4. posiada deklarację klucza głównego (jedna lub wiele kolumn)

SQLAlchemy - analogia do SQL

Postać SQL

```
SELECT * FROM klienci;
```

Postać ORM

```
Klient.query.all()
```



```
SELECT * FROM klienci
```

```
WHERE age > 40;
```



```
Klient.query.filter(age > 40)
```

SQLAlchemy - analogia do SQL

Postać SQL

```
SELECT COUNT(*) FROM  
klienci;
```

Postać ORM

```
Klient.query.count()
```



Więcej informacji nt filtrowania:

<https://docs.sqlalchemy.org/en/latest/orm/tutorial.html#common-filter-operators>

Kilka trików, które warto znać

Triki

Wskazanie bazy danych:

- `engine = create_engine('driver://connection_string')`

np:

- `en = create_engine('sqlite:///f.db')` – ścieżka względna(!)
- `en = create_engine('sqlite://')` – baza wyłącznie w RAM
(najszybsza 😊 i najbardziej ulotna ☹)

Triki

Włączenie logowania poleceń:

- `en = create_engine('sqlite:///foo.db', echo = True)`

```
2018-12-10 22:15:10,107 INFO sqlalchemy.engine.base.Engine SELECT CAST('test plain returns' AS VARCHAR(60)) AS anon_1
2018-12-10 22:15:10,107 INFO sqlalchemy.engine.base.Engine ()
2018-12-10 22:15:10,109 INFO sqlalchemy.engine.base.Engine
CREATE TABLE portale (
    idx INTEGER NOT NULL,
    nazwa_portalu VARCHAR,
    PRIMARY KEY (idx)
)
2018-12-10 22:15:10,109 INFO sqlalchemy.engine.base.Engine ()
2018-12-10 22:15:10,117 INFO sqlalchemy.engine.base.Engine COMMIT
```

Triki

Wyświetlenie SQL odpowiadającego zapytaniu:

```
>> max_val = session.query(func.max(Goods.price)).one()  
>> print('Max price:', max_val[0])  
4
```

Nie ma znaczenia, czy to one(), first(), limit()

```
>> query = session.query(func.max(Goods.price))           # pozbawione one()  
>> print(query)  
  
SELECT max(goods.price) AS max_1  
FROM goods
```

Triki

Pytanie:

- Jak często należy wykonać commit() przy serii operacji insert?

Pytanie:

- Co jeśli wyników zapytania jest więcej niż mamy pamięci RAM?

~~query = session.query(Goods).all()~~ - ryzykowne podejście

Lepsze podejście:

```
for element in session.query(Goods):      # to jest iterator
    print(element)                        # o ile mamy metodę __repr__() w klasie Goods
```


Triki

- Czy mogę pominąć `and_()`?

Tak, jeśli to jedyny łącznik warunków. Wtedy to są konstrukcje równoważne:

```
session.query(Goods).filter(and_(Goods.origin == 'Poland', Goods.price>0))
```

```
session.query(Goods).filter(Goods.origin == 'Poland', Goods.price>0)
```

- Czy mogę nadać kolumnie nazwę `id`?

Możesz, ale przykrywasz funkcję wbudowaną `id()`. Lepsze nazwy dla kolumny z kluczem głównym to: `idx`, `nazwa_tabeli_id`, `pk`.

Triki (już nie SQLAlchemy)

- Jak wyłączyć indeksowanie plików w konkretnym katalogu w PyCharm, a mam tam 2500 plików html po sesji webscrappingu?

File -> Settings -> Editor -> File Types -> Ignore files and folders

- Pamiętasz do czego służy setattr()?

Może przydać się do uzupełnienia atrybutów obiektu, gdy tych atrybutów jest relatywnie dużo i jednolinijkowiec przypomina stonogę ;-)

setattr(obiekt, „atrybut jako string”, wartość)

Omówienie rozwiązań zadania domowego

Wasze najciekawsze doświadczenia

Dzięki!

