

Bestudeer de video van de volgende les geheel:

Object modellering: van stabiel object model naar flexibele code

Experimenteer zelfstandig met de daarin behandelde voorbeeld-programma's. Een link naar de broncode daarvan staat onder de video. In de volgende les kun je hierover vragen stellen.

Opgaven bij de huidige les:

In te leveren uiterlijk 2 dagen vóór de volgende les.

Niveau 1

Installeer de packages *requests* en *beautifulsoup*.

Lees een tutorial over het gebruik van deze packages, bijvoorbeeld:

<https://www.digitalocean.com/community/tutorials/how-to-work-with-web-data-using-requests-and-beautiful-soup-with-python-3>

Schrijf een webscraping-programma dat van de webpagina *www.wiztech.nl* eerst alle koppen op niveau 1 t/m 3 afdruckt en daarna de structuur van de hele webpage. Dit laatste is de zogenaamde DOM-tree. DOM komt van Domain Object Model, onderwerp van de volgende les.

Niveau 2

Maak een programma dat vraagt een term in te typen en die term vervolgens uitlegt. De verzameling termen is uit te breiden door modules toe te voegen. Elk onderwerp heeft een aparte module met daarin een verzameling termen rond een bepaald onderwerp, bijvoorbeeld techniek, muziek of sport, en bijbehorende uitleg. Indien een term niet bekend is, wordt dit aan de gebruiker meegedeeld. Indien een term op meerdere plaatsen voorkomt, worden alle betekenissen gegeven.

Extra: Zorg dat je om kunt schakelen tussen nederlandstalige en engelstalige termen en uitleg.

Niveau 3

Dit is een combinatie van niveau 1 en 2. Maak een module die aan de hand van een lijst van URL's de inhoud van webpagina's ophaalt en ontleedt. Maak een andere module die alle ASCII of Unicode tekst bestanden uit een bepaalde lokale directory leest. Maak een programma dat beide modules combineert om naar plagiaat te zoeken.

N.B. Niet elk klein woordgroepje dat je tegenkomt is plagiaat, probeer een zinnig, instelbaar criterium te gebruiken. In hoeverre moeten teksten overeenkomen om van plagiaat te spreken en hoe meet je dat? Zorg dat ook een bestand met uitzonderingen wordt geladen. De zinsnede "en ze leefden nog lang en gelukkig" is wat dit betreft bijvoorbeeld betrekkelijk onschuldig.

N.B. 2. Het maken van een praktisch bruikbaar tool gaat wat te ver voor deze lessen. De bedoeling is, dat je laat zien dat je met Python, packages en modules uit de voeten kunt!

--//--