

ALGORITMA DAN STRUKTUR DATA
MODUL 7
REGULAR EXPRESSIONS



Disusun oleh:
Muhammad Ferizal Fadhli
L200210119
D

TEKNIK INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA
2022/2023

Soal Soal Mahasiswa

1. Meng-ekstrak kata-kata dengan awalan 'me'. Bukalah halaman web id.wikipedia.org/wiki/Indonesia. Salin seluruh tubuh artikel itu dan tuangkan ke dalam notepad. Simpan sebahai Indonesia.txt. Tugasmu adalah mengekstrak kata yang mempunyai awalan 'me', lalu menyimpannya dalam sbuah tuple. Perhatikan bahwa 'me' bisa muncul di awal tengah ataupun akhir kalimat.

Berikut Kode Program :

```
import re

isi = ()
f = open("Indonesia.txt", "r", encoding="latin1")
teks = f.read()
p = r"me\\w+"
string = re.findall(p, teks)
p2 = r"\\w+me\\w+"
string2 = re.findall(p2, teks)
p3 = r"\\w+me"
string3 = re.findall(p3, teks)
isi = string + string2 + string3
print(isi)
```

Berikut SS Ketika Program Dijalankan :

```
ms-python.python:2023.8.0\python_files\lib\python\debugger\adapter... 36244 ... "D:\Vagus Kuliah\Semester 4\Praktikum ASO\Portman\Ano1.py",
"merupakan", "menjadi", "mencapai", "merika", "merintahan", "merintah", "menetapkan", "menepati", "mengagungkan", "merjadi", "menjalin", "membawa", "memonopoli", "memproklamasikan", "merdekaan", "mendapat", "mendia",
"membentuk", "mencapai", "membentuk", "memiliki", "memiliki", "mendukung", "merupakan", "Amerika", "pemerintahan", "pemerintah", "kemerdekaan", "Am", "pew", "pew", "Hinduisme", "Buddhisme", "Hinduisme", "Buddhisme",
"kom", "senaraiisme"]
```

2. Meng-ekstrak kata-kata dengan awalan 'di'. Sama seperti di atas, tapi yang kamu ekstrak adalah kata kata dengan awalan 'di'. Perhatikan yang dimaksud disini adalah awalan 'di' (menuntukkan kata kerja pasif), bukan kata depan 'di' (menunjukkan keterangan tempat), yang kita bahas pada nomer dibawah.

Berikut Kode Program :

```
import re

isi = ()
f = open("Indonesia.txt", "r", encoding="latin1")
teks = f.read()
p = r"di\\w+"
string = re.findall(p, teks)
isi = string
print(isi)
```

Berikut SS Ketika Program Dijalankan :

```

nlp_python.py[nlp_python-2022-07-07pycnlp-215555pycnlp-udagpy[udagpy]-7.7.7[udagpy[udagpy]-2022-07-07]] - D:\ragas_kuliah\Semester 4\Praktikum ASJP\Kerendian\7\udag.py
['dikenal', 'dilintasi', 'dikenal', 'dia', 'dipakai', 'disebut', 'dia', 'dicalonkan', 'dipilih', 'dipengaruhi', 'dirinya', 'dia', 'dia', 'diri', 'diri', 'diami']
PS D:\Tugas Kuliah\Semester 4\Praktikum ASJP\Kerendian\7>

```

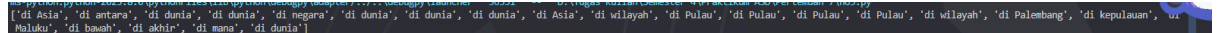
3. Meng-ekstrak kata depan ‘di’ dan tempat yang ditunjukkan. Di sini tugasmu adalah mengekstrak masih pada berkas Indonesia.txt – kata depan ‘di’ beserta satu kata yang mengikutinya.

Berikut Kode Program :

```
import re

isi = ()
f = open("Indonesia.txt", "r", encoding="latin1")
teks = f.read()
p = r"di \w+"
string = re.findall(p, teks)
isi = string
print(isi)
```

Berikut SS Ketika Program Dijalankan :



4. Mengekstrak data dari berkas html. Tugamu di sini adalah membuat sebuah list of tuples yang memuat nama-nama negara innovation index-nya. Kunjungilah halaman en.wikipedia.org/wiki/Knowledge_Economic_index. Perhatikan halaman itu memuat sebuah besar hasil penelitian.

Berikut Kode Program :

```
import re

f= open("KEI.html", "r", encoding="latin1")
teks = f.read()

p = r'([\w+]+)</a></td>\n<td>([0-9.]+)</td>' #mencari pola
string = re.findall(p, teks) #mencari string
list = []
for i in string:
    a = (i[0], float(i[1]))
    list.append(a)
print(list)
```

Berikut SS Ketika Program Dijalankan :

