

# Μεταγλωττιστές 2020

## Προγραμματιστική Εργασία #2

**Ονοματεπώνυμο:** Σωκράτης Καραουσάς

**A.M.:** Π2017033

### 1ο Ερώτημα

`('<title>(.*?)</title>')`

Η χρήση της τελείας (.) και της πρόσθεσης (+) είναι για το ταίριασμα οποιδήποτε κειμένου ανάμεσα στα `<title>` και `</title>`.

### 2ο Ερώτημα

`('<!--.*?-->',re.DOTALL)`

Η χρήση της τελείας (.) και του αστεριού (\*) είναι για το ταίριασμα οποιδήποτε σχόλιου όπου βρίσκεται ανάμεσα στα `!<--` και `-->` και δεν χρειάζεται να υπάρχει έστω ένας χαρακτήρας λόγω του \* αντί του +.

### 3ο Ερώτημα

`(r'<(s(?:cript|tyle)).*?>.*?</\1>',re.DOTALL)`

Η χρήση του `(s(?:cript|tyle))` είναι για το ταίριασμα των tags `<script>` και `</script>` και `<style>` και `</style>`. Η χρήση του ερωτηματικού (?) είναι για να μην οριστεί νέο group. Στόχος της έκφρασης αυτής, είναι η αποθήκευση κάθε ονόματος του tag μέσα στο `group(1)`.

### 4ο Ερώτημα

`(r'<a.+?href="(.*?)".*?>(.*?)</a>',re.DOTALL)`

Η έκφραση αυτή είναι για το ταίριασμα της href και το περιεχόμενο των `<a>` και `</a>`.

### 5ο Ερώτημα

Δύο εκφράσεις:

1. `(r'<.+?>|</.+?>',re.DOTALL)`

2. `(r'<.+?/>',re.DOTALL)`

Ο ρόλος αυτών των εκφράσεων, αποτελεί το ταίριασμα με κάθε html tag του κειμένου

### 6ο Ερώτημα

`(r'&(amp|gt|lt|nbsp);')`

Εδώ ο ρόλος της έκφρασης, αποτελεί το ταίριασμα όλων των &amp, &gt, &lt και &nbsp του κειμένου.

### **7ο Ερώτημα**

(r'\s+')

Εδώ ο ρόλος της έκφρασης είναι το ταίριασμα όλων των χαρακτήρων whitespace.