



DIS08 – Data Modeling

04 – Regex – Tutorial Session

Fabian Haak, Technische Hochschule Köln, Cologne, Germany

Version: 2021-11-11

Topics for Today

1. Q&A

2. Quiz

3. Assignment/ Help

Open Questions



Topics for Today

1. Q&A

2. Quiz

3. Assignment/ Help

Quiz!

Das heutige Quiz lösen wir gemeinsam. Sie werden hier Aufgaben gestellt bekommen, für die Sie jeweils ca 3 Minuten Für die Bearbeitung haben, anschließend tragen wir die Lösungen Zusammen.

Das Quiz basiert auf regexone.com. Es ist empfohlen, dort das interaktive Tutorial abzuschließen.

Mit Tools wie auf regex101.com können Sie Ihre RegEx testen.

Warmup 1: ABBA

Welcher der folgenden Terme wird nicht von der RegEx $(AB|(BA)^*)^*$ gematched?

1. ABBA
2. ABABABAB
3. ABBAAB
4. ABBAABAA

Warmup 2: Hallo Welt

Welcher der folgenden Strings wird nicht komplett von $((\backslash w|\backslash s))^*$ gematcht?

1. Hallo Welt
2. hallo welt
3. hallo welt!
4. HALLO WELT

Warmup 3: Batman???

Welcher der folgenden Strings wird nicht komplett von
`((na){4}\s?)*batman\!?`
Gematcht? Warum?

1. nananananananana batman!
2. nananana batman
3. nananananananabatman!
4. nananana nananana batman!
5. batman!

Question 1: Es war einmal...

Schreiben Sie eine RegEx, die alle Inputs komplett matched, Die mit "Es war einmal" oder "Es begab sich" anfangen.

Question 2: Matching Decimal Numbers

Für die Aufgabe “Problem 1” wird auf Regexone.com die folgende Lösung gegeben:

`^-?\d+(, \d+)*(\. \d+(e\d+)?)?$`

Die RegEx soll alle Zahlen in Amerikanischer Notation (also mit einem Komma als Tausendertrennzeichen und einem Punkt zum Abtrennen der Nachkommastellen) matchen:

Match	-255.34	✓
Match	128	✓
Match	1.9e10	✓
Match	123,340.00	✓
Skip	720p	

`^-?\d+(, \d+)*(\. \d+(e\d+)?)?$`
Continue ›

Die gegebene Lösung hat jedoch (mindestens) einen Fehler! Finden sie einen String, der gematcht wird, aber keine korrekt notierte Zahl ist!

Question 3: Matching Filenames

Schreiben Sie eine RegEx, die Dateien mit den Extensions .jpg, .gif und .png Dateien sowie alle Dateien, die den Term "log" VOR der Extension enthalten.

Question 4: Email-Adressen überprüfen

Szenario: Sie möchten für eine Website eine regular Expression schreiben, mit der Sie überprüfen können, ob ein eingegebener Text eine Email-Adresse ist.

Es sollen folgende Kriterien geprüft werden:

1. am Anfang muss der so genannte local-part stehen, der nur aus Buchstaben (klein und groß), Ziffern, Punkten, Bindestrichen und/oder Unterstrichen bestehen darf und mindestens 5, höchstens 30 Zeichen enthalten muss
2. es muss ein @-Zeichen folgen.
3. es folgt der Domain-part, dessen erster Teil nur Buchstaben, Ziffern und Bindestriche enthalten darf und zwischen 2 und 20 Zeichen lang sein darf
4. Der Domain-part muss auf einen Punkt gefolgt von zwei bis drei Buchstaben enden (z.B. ".de" oder ".com")

Auf der folgenden Folie finden Sie einige Testadressen.

Question 4: Teststrings

Folgende Strings können Sie zum Testen verwenden, die ersten 4 sollen erfolgreich vollständig gematched werden, die letzten 3 dürfen nicht akzeptiert werden.

Fabian.Haak@TH-Koeln.de
fake_Adresse.de@me.com
Walter-Mueller1962@KrankenkasseSaarland.fr
12345@678.de

Trink/.dich/.fit@koeln.de
123@456.78
Käsebüro.Kassel@t-online.de

Topics for Today

1. Q&A

2. Quiz

2. Assignment/ Help