

Úvod do spracovania prirodzeného jazyka

Matúš Pikuliak

Motivácia #1

Každú sekundu (v roku 2016)

- 6,000 tweetov
- 40,000 google searchov
- 2,000,000 emailov

Na internete je (v roku 2013) **14.5 biliónov** stránok

PubMed obsahuje **14 miliónov** článkov o medicíne

Motivácia #2



Aplikácie NLP - Preklad

Translate



English Spanish French English - detected ▾



English Spanish Scots Gaelic ▾

Translate

Hello World



Hello an t-Saoghail






Suggest an edit

See also

World

Aplikácie NLP - Vyhľadávanie, sumarizácia


Google   [Sign in](#)

[All](#) [Images](#) [Shopping](#) [News](#) [Videos](#) [More ▾](#) [Search tools](#) 

About 2,090,000 results (0.42 seconds)

Images for Grape soda

[Report images](#)



[More images for Grape soda](#)

Urban Dictionary: grape soda

www.urbandictionary.com/define.php?term=grape+soda ▾ Urban Dictionary ▾


(1) Miles arrived to the anticipated enthusiasm of his ghetto brethren after picking up a sixer of grape soda from the convenience store. (2) The neighborhood ...

Grape soft drink - Wikipedia, the free encyclopedia

https://en.wikipedia.org/wiki/Grape_soft_drink ▾ Wikipedia ▾

Grape soft drinks (also known as grape drink, grape soda, or grape pop in certain regions of the US) are typically sweetened carbonated drinks with a grape ...


Grape soft drink



Grape soft drinks are typically sweetened carbonated drinks with a grape flavor. Grapeade first appeared as a variety of carbonated drink provided in soda fountains in American drugstores in the late ...
[Wikipedia](#)

People also search for

[View 5+ more](#)



[Orange drink](#) [Cream soda](#) [Root beer](#) [Kool-Aid](#) [Enhanced water](#)

[Feedback](#)

See results about


Aplikácie NLP - Objavovanie znalostí

Google

[Web](#) [Shopping](#) [Images](#) [News](#) [Videos](#) [More ▾](#) [Search tools](#)

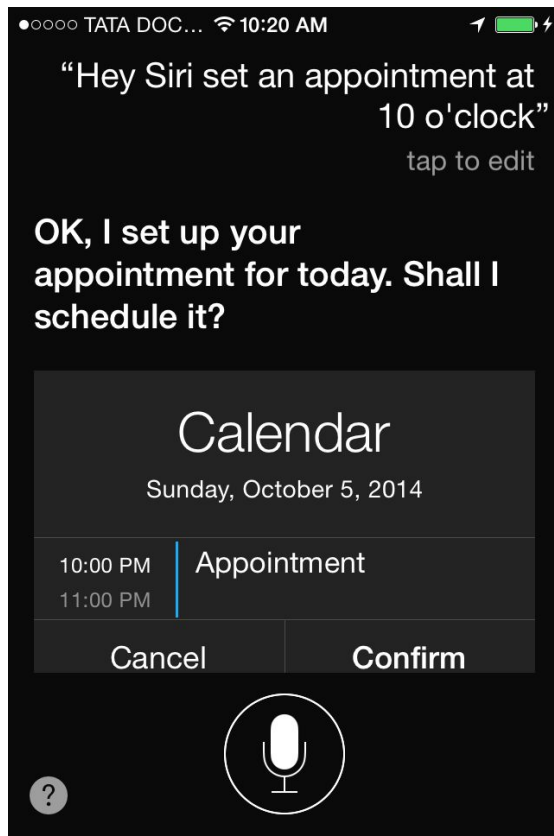
About 13,100 results (0.49 seconds)

Alan Alda
Played Hawkeye Pierce on M*A*S*H



[Get updates about Alan Alda](#)

Aplikácie NLP - Osobní asistenti



Aplikácie NLP - Ďalšie

- klasifikácia textov (spam, jazyk, sentiment)
- anotácia slov (vetný člen, význam homoným, výslovnosť)
- rozpoznávanie sémantiky
- generovanie textu (sumarizácia, titulkovanie obrázkov, preklad)
- speech-to-text, text-to-speech, rozpoznávanie a imitácia písma, spracovanie znakového jazyka
- detekcia a opravovanie gramatických chýb

Metódy - #1 štatistické

- pracujeme s veľkými korpusmi
- používame metódy strojového učenia, v súčasnosti najmä NN
- pozeráme sa na to, aké slová sa štatisticky vyskytujú pri sebe a pod.

Metódy - #2 syntaktické

- pracujeme s vetami
- vetu sa snažíme rozbiť na syntaktický strom a vyčítať z neho jej význam
- pozeráme sa na to, aký vzťah majú slová vo vete

Vrstvy jazyka

1. fonetika, fonológia, morfológia - úroveň slov
2. syntax - úroveň skladania viet
3. sémantika - úroveň významu viet
4. pragmatika - úroveň použitia jazyka v praxi

<http://text.fiit.stuba.sk>