

Vývin extrakcie informácií

Metódy inžinierskej práce 2023/2024

Nikolas Kristín

Fakulta informatiky a informačných technológií
Slovenská technická univerzita v Bratislave

26. november 2023

Ešte pre prehľadom prezentácie sa zvyčajne uvádza motivácia.
Bežne sa aj tu používajú odrážky, ale tento text je naschvál uvedený bez odrážok. Niekedy môže byť potrebné uviesť aj citáty. . .
Toto je len príklad slajdov. Ako urobiť dobrú prezentáciu bolo vysvetlené na prednáške.

Prehľad

1 MUC

2 Metódy extrakcie informácií

- Regex
- Named Entity Recognition
- Relation Extraction
- Customized Data Pipelines

3 Extrakčné systémy

- Jasper
- Rosoka

Message Understanding Conference (MUC)

<i>Conference</i>	<i>Year</i>	<i>Text Source</i>	<i>Topic (Domain)</i>
MUC-1	1987	Mil. reports	Fleet Operations
MUC-2	1989	Mil. reports	Fleet Operations
MUC-3	1991	News reports	Terrorist activities in Latin America
MUC-4	1992	News reports	Terrorist activities in Latin America
MUC-5	1993	News reports	Corporate Joint Ventures, Microelectronic production
MUC-6	1995	News reports	Negotiation of Labor Disputes and Corporate Management Succession
MUC-7	1997	News reports	Airplane crashes, and Rocket/Missile Launches

Informácie jednotlivých konferencií, Zdroj: en.wikipedia.org/wiki/MessageUnderstandingConference

Regex

Named Entity Recognition

Relation Extraction

Customized Data Pipelines

Jasper

Rosoka