

ZADANIE BAKALÁRSKEHO PROJEKTU

Meno študenta: **Cimerman Matúš**
Študijný odbor: Informatika
Študijný program: Informatika
Názov projektu: **Analýza prúdu údajov**

Zadanie:

V ostatnom čase sa štandardom pre spracovanie veľkého objemu údajov stalo distribuované spracovanie pomocou map-reduce paradigmy. Rôzne nástroje založené na tejto paradigme sú vhodné pre dávkové spracovanie veľkého objemu údajov, ktoré však nedokážu poskytnúť presný pohľad na údaje meniace sa v reálnom čase. Pre spracovanie takýchto prúdov údajov vzniklo viacero prístupov (napr. Lambda Architektúra) a nástrojov (napr. Storm), ktoré sa nepozerajú na spracovávané údaje ako na statickú kolekciu, ale ako na potenciálne neobmedzený prúd nových údajov.

Analyzujte súčasný stav spracovania prúdov údajov a rôzne nástroje a techniky používané pri spracovaní takýchto prúdov. Zvoľte si jeden z existujúcich prístupov alebo nástrojov a overte jeho vlastnosti na rôznych úlohách spracovania prúdu údajov ako je napríklad analýza trendov v prúde správ zo služby Twitter alebo v prúde údajov z populárneho skracovača URL adres Bitly.

Práca musí obsahovať:

Anotáciu v slovenskom a anglickom jazyku
Analýzu problému
Opis riešenia
Zhodnotenie
Technickú dokumentáciu
Zoznam použitej literatúry
Elektronické médium obsahujúce vytvorený produkt spolu s dokumentáciou

Miesto vypracovania: Ústav informatiky a softvérového inžinierstva, FIIT STU, Bratislava
Vedúci projektu: Ing. Jakub Ševcech

Termín odovzdania práce v zimnom semestri: 10. 12. 2014
Termín odovzdania práce v letnom semestri: 12. 5. 2015



Bratislava 22. 9. 2014

prof. Ing. Pavol Návrat, PhD.
riaditeľ ÚISI