

# Techniky spracovanie veľkých dát\*

Tomáš Zenka

Slovenská technická univerzita v Bratislave  
Fakulta informatiky a informačných technológií  
xzenka@stuba.sk

5. november 2023

## Abstrakt

Článok skúma a porovnáva techniky spracovania veľkého množstva dát, čo je kľúčové v dnešnej digitálnej dobe, kde sa generuje obrovské množstvo informácií. Cieľom je poskytnúť prehľad o moderných prístupoch a nástrojoch určených na manipuláciu s masívnymi dátovými súbormi. Tieto nástroje zahŕňajú distribuované systémy na spracovanie dát, algoritmy strojového učenia a metriky na hodnotenie kvality dát. Dôraz sa kladie na potrebu rýchleho spracovania dát v reálnom čase, čo umožňuje rýchlu analýzu a tvorbu hodnotných poznatkov z týchto objemných dátových zdrojov. Článok taktiež uvádza príklady aplikácií v rôznych odvetviach, ako je medicína, finančníctvo a priemysel. Spracovanie veľkého množstva dát sa stáva nevyhnutným nástrojom pre konkurencieschopnosť a inovácie v súčasnom digitálnom prostredí.

## 1 Úvod

Zatiaľ len v odrážkach

- úvod do problematiky
- definícia veľkých dát
- nárast objemu dát
- výzvy spojené so spracovaním veľkých dát
- cieľ a štruktúra článku

## 2 Základné techniky spracovania veľkých dát

Stručný popis (1 - 2 vety) na techniky, ktoré môžeme použiť.

1. NoSQL databázy 2.1

2. In-memory processing 2.2

---

\*Semestrálny projekt v predmete Metódy inžinierskej práce, ak. rok 2023/24, vedenie: Vladimír Mlynarovič

## 2.1 NoSQL databázy

## 2.2 In-memory processing

## 3 Iná časť

Môže sa zdať, že problém vlastne nejestvuje [1], ale bolo dokázané, že to tak nie je [2, 3]. Napriek tomu, aj dnes na webe narazíme na všelijaké pochybné názory [4]. Dôležité veci možno *zdôrazniť kurzívou*.

### 3.1 Nejaké vysvetlenie

Niekedy treba uviesť zoznam:

- jedna vec
- druhá vec
  - x
  - y

Ten istý zoznam, len číslovaný:

1. jedna vec
2. druhá vec
  - (a) x
  - (b) y

### 3.2 Ešte nejaké vysvetlenie

**Veľmi dôležitá poznámka.** Niekedy je potrebné nadpisom označiť odsek. Text pokračuje hneď za nadpisom.

## 4 Dôležitá časť

## 5 Ešte dôležitejšia časť

## 6 Záver

## Literatúra

- [1] James O. Coplien. *Multi-Paradigm Design for C++*. Addison-Wesley, 1999.
- [2] Krzysztof Czarnecki, Simon Helsen, and Ulrich Eisenecker. Staged configuration through specialization and multi-level configuration of feature models. *Software Process: Improvement and Practice*, 10:143–169, April/June 2005.
- [3] Krzysztof Czarnecki and Chang Hwan Peter Kim. Cardinality-based feature modeling and constraints: A progress report. In *International Workshop on Software Factories, OOPSLA 2005*, San Diego, USA, October 2005.

- [4] Carnegie Mellon University Software Engineering Institute. A framework for software product line practice—version 5.0. [http://www.sei.cmu.edu/productlines/frame\\_report/](http://www.sei.cmu.edu/productlines/frame_report/).