

# Fulltextové vyhľadávanie so štruktúrovaným výsledkom

Vedúci práce: RNDr. Peter Gurský, PhD.

Autor: Patrik Sedlák



# Motivácia k výberu témy

- Riešenie problému v praktickej oblasti
- Projekt, na ktorom pracuje celý tím ľudí
- Možnosť spoznať nové technológie
- Práca v oblasti webu



# Prečo fulltextové vyhľadávanie ?

- Pri fulltextovom vyhľadávaní sa hľadá akákoľvek zhoda v dokumente s vyhľadávaným výrazom
- Fulltextové vyhľadávanie môže používateľovi ponúknuť širšiu paletu výsledkov
- Často však veľa z týchto výsledkov nie je vyhovujúcich – sú vytrhnuté z kontextu



# Prečo fulltextové vyhľadávanie ?

- Pri fulltextovom vyhľadávaní sa hľadá akákoľvek zhoda v dokumente s vyhľadávaným výrazom
- Fulltextové vyhľadávanie môže používateľovi ponúknuť širšiu paletu výsledkov
- Často však veľa z týchto výsledkov nie je vyhovujúcich – sú vytrhnuté z kontextu

Čo je problém



# VYHĽADÁVANIE: SAMSUNG 7"

Nájdene v produktoch (208 výsledkov)

Nájdene v stránkach (24 stránok)

Nájdene v článkoch (3 články)

**i** Pre hľadaný výraz Samsung 7" existujú aj aktuálne vypredané produkty. [Zobraziť aj tieto produkty](#) alebo [zobraziť len tieto produkty](#)



**Samsung Galaxy S7 (biely)**

- Uhlopriečka displeja: 5.1"
- Poskytovaný OS: Android
- Vnútna kapacita pre ukladanie: 32 GB

5.1"



**Samsung Galaxy S7 edge (čierny)**

- Uhlopriečka displeja: 5.5"
- Poskytovaný OS: Android
- Vnútna kapacita pre ukladanie: 32 GB

5.5"



**Samsung Galaxy S7 edge (biely)**

- Uhlopriečka displeja: 5.5"
- Poskytovaný OS: Android
- Vnútna kapacita pre ukladanie: 32 GB

5.5"



**Samsung Galaxy S7 (zltý)**

- Uhlopriečka displeja: 5.1"
- Poskytovaný OS: Android
- Vnútna kapacita pre ukladanie: 32 GB

5.1"

# VYHĽADÁVANIE: SAMSUNG 7"

Nájdene v produktoch (208 výsledkov)

Nájdene v stránkach (24 stránok)

Nájdene v článkoch (3 články)

**i** Pre hľadaný výraz Samsung 7" existujú aj aktuálne vypredané produkty. [Zobraziť aj tieto produkty](#) alebo [zobraziť len tieto produkty](#)



Samsung Galaxy S7 (biely)

- Uhlopriečka displeja: 5.1"
- Poskytovaný OS: Android
- Vnútrotná kapacita pre ukladanie: 32 GB

5.1"



Samsung Galaxy S7 edge (čierny)

- Uhlopriečka displeja: 5.5"
- Poskytovaný OS: Android
- Vnútrotná kapacita pre ukladanie: 32 GB

5.5"



Samsung Galaxy S7 edge (biely)

- Uhlopriečka displeja: 5.5"
- Poskytovaný OS: Android
- Vnútrotná kapacita pre ukladanie: 32 GB

5.5"



Samsung Galaxy S7 (zltý)

- Uhlopriečka displeja: 5.1"
- Poskytovaný OS: Android
- Vnútrotná kapacita pre ukladanie: 32 GB

5.1"

Kategória?

Heureka.sk » **Vyhľadávanie Samsung 7"**

## 🔍 Nájdené kategórie (20)

[Ochranné fólie pre tablety \(116\)](#), [Puzdrá na tablet PC a čítačky ekníh \(111\)](#), [Foto - Video batérie - neoriginálne \(95\)](#), [LCD monitory \(56\)](#), [Tablety \(46\)](#), [Tvrdené skla pre mobilné telefóny \(44\)](#), [Stylusy \(15\)](#), [Náplne a tonery - originálne \(7\)](#), [Gamepady \(7\)](#), [Náplne a tonery - kompatibilné \(6\)](#), [Viac kategórií »](#)

Cena (€) Od:  Do:

Tovar  Nové  bazárové

Zoradiť podľa  ▼

Zobrazujem 1 - 20 zo 590 produktov [Ďalšie »](#)



### [Samsung Galaxy S7 Edge G935F 32GB](#)

(Mobilné telefóny) ★★★★★ (24 recenzií)

Vyvážený design kráčajúci ruka v ruke s dokonalou funkčnosťou? To je nový model Galaxy Edge, ktorého špičková výbava posúva hranice inteligentných telefónov opäť o kúsok ďalej.

[Uložiť na porovnanie](#)

**679,00 - 953,78 €**  
v [67 obchodoch](#)

[Porovnať ceny](#)



### [Samsung Galaxy A3 2016 A310F](#)

(Mobilné telefóny) ★★★★★ (11 recenzií)

Tento vdarený smartphone spoločnosti Samsung ponúka dokonalý vzhľad, vysoký výkon a výborný fotoaparát, s ktorým vyhotovíte perfektné zábery aj v zlých svetelných podmienkach.

[Uložiť na porovnanie](#)

**219,99 - 355,06 €**  
v [70 obchodoch](#)

[Porovnať ceny](#)

Kategórie 😊

7 palcov ? ☹️

# Krok za krokom

- Naštudovať Wicket framework
- Nainštalovať Elasticsearch a naučiť sa s ním pracovať
- Naštudovať literatúru k danej tematike
- Navrhnuť respektíve modifikovať už existujúcu metódu fulltextového vyhľadávania
- Implementovať metódu v projekte Kapsa
- Otestovať a analyzovať navrhnutú metódu
- Skoncipovať a formulovať text k práci





# Ako fulltextovo vyhľadávať ?

Využijeme Elasticsearch, ktorý:

- vyhľadáva nad JSON štruktúrovanými dokumentmi pričom výsledok je opäť štruktúrovaný JSON dokument
- je zdarma
- poskytuje fulltextové vyhľadávanie
- dáta v reálnom čase
- vysoká dostupnosť

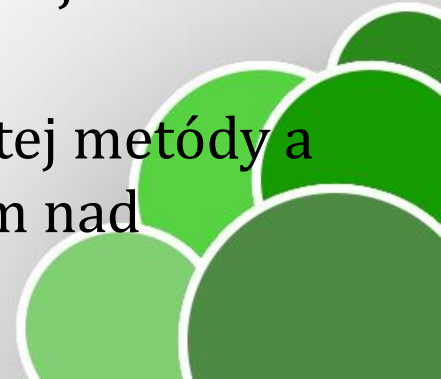


elasticsearch.



# Ciele práce

- Navrhnuť metódu fulltextového vyhľadávania nad produktmi internetových obchodov, ktorej výsledkom je okrem zoznamu produktov, ktoré najviac zodpovedajú vstupnej otázke, aj štruktúrovaná informácia o úspešnosti vyhľadávania podľa domén a atribútov.
- Implementovať navrhnutú metódu ako modul v projekte kapsa.
- Analyzovať rýchlosť, presnosť a úplnosť navrhnutej metódy a jej porovnanie bežným fulltextovým vyhľadávaním nad reálnymi dátami.



# Odporúčaná literatúra

- [1] Carstens, Carola ; Womser-Hacker, Christ: Ontology Based Query Expansion - Retrieval Support for the Domain of Educational Research. 2012 [online] Dostupné na <<http://d-nb.info/1023809400>>.
- [2] Baeza-Yates, Ricardo A. and Ribeiro-Neto, Berthier: Modern Information Retrieval, 1999 ISBN 020139829
- [3] Tvarožek, M., Bieliková, M. Bridging Semantic and Legacy Web Exploration: Orientation, Revisitation and Result Exploration Support. In: 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. [online] Dostupné na <<http://acadmedia.wku.edu/Zhuhadar/WI-IAT%202010/Data/4191c326.pdf>>
- [4] Jozef Kuper: Unifikácia metadát o produktoch internetových obchodov. Diplomová práca, PF UPJŠ, 2015





Ďakujem za pozornosť