

Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Katedra informačních technologií

Studijní program: Aplikovaná informatika
Studijní obor: Informační systémy a technologie

VYHLEDÁVÁNÍ JAKO SLUŽBA

DIPLOMOVÁ PRÁCE

Student: Bc. Luděk Veselý
Vedoucí: Prof. Ing. Zdeněk Molnár, CSc.

2017

University of Economics in Prague
Faculty of Informatics and Statistics
Department of Information Technologies

Study programme: Applied Informatics
Study branch: Information Technologies

SEARCH AS A SERVICE

DIPLOMA THESIS

Student: Bc. Luděk Veselý
Supervisor: Prof. Ing. Zdeněk Molnár, CSc.

2017

Tento list nahradte
originálem zadání.

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracoval samostatně a že jsem uvedl všechny použité prameny a literaturu, ze které jsem čerpal.

Datum:

Podpis:

Abstrakt

Tato diplomová práce popisuje návrh a tvorbu fulltextového vyhledávání poskytovaného jako služba.

Klíčová slova

Fulltext, Elasticsearch

Abstract

This diploma thesis describes creation of fulltext search service.

Key words

Fulltext, Elasticsearch

Poděkování

Rád bych poděkoval Ing. Ivanovi Jelínkovi za rady a pomoc při řešení.

Obsah

1	Úvod	8
1.1	Cíle práce	8
1.2	Cílová skupina	8
1.3	Použité metody	8
1.4	Struktura práce	8
2	Komentovaná rešerše informačních zdrojů	9
3	Samotná práce...	10
3.1	Analyza problemu a požadavku na aplikaci	10
3.2	Fulltextové hledání obecně	10
3.3	Indexace	10
3.4	Nastroje použitelné pro fulltextové vyhledávání	10
3.5	Architektura aplikace	10
3.6	Testování, overení	11
4	Závěr	12
4.1	Dosažení vytčených dílů	12
A	Obsah přiloženého DVD	15
B	Graf jako příloha	16

1 Úvod

V této diplomové práci práci...

1.1 Cíle práce

...

1.2 Cílová skupina

...

1.3 Použité metody

...

1.4 Struktura práce

...

2 Komentovaná rešerše informačních zdrojů

...

3 Samotna prace...

3.1 Analýza problému a požadavku na aplikaci

Vysvětlení problému, který je řešen Specifika prostředí e-commerce, definice kontextu, ve kterém bude nástroj používán Definice požadavku na samotné vyhledávání, akceptační kritéria Technické požadavky (datové formáty, rychlost odezvy...) Popis oborů, kterých se práce týká (získávání informací, NLP, big data...) Analýza problému fulltextového vyhledávání

3.2 Fulltextové hledání obecně

3.3 Indexace

Popis problému indexování dokumentu Analýza dat použitých k vyhledávání Zpracování přirozeného jazyka

Sklonování, časování, tvaroslovi obecně Vaha slov, stop slova Preklepy Synonyma, zargón Lemmatizace Tezaurus Navrh řešení

3.4 Nástroje použitelné pro fulltextové vyhledávání

MySQL Elasticsearch Sphinx Algolia Navrh architektury aplikace

Vyber technologie a nástroje Definice případu užití Navrh architektury aplikace Implementace aplikace

3.5 Architektura aplikace

Popis služeb, modulu, vztahu mezi nimi Database

Nastavení a nasazení Elasticsearch Implementace indexace Implementace vyhledávání Backend

Popis tríd, rozhraní, implementační detaily Frontend

Navrh, komunikace s backendem Deployment

CI, nasazování na server

3.6 Testovani, overeni

Overeni funkcnosti a prinosu Testovani, zda jsou splnena akceptacni kriteria Porovnani s stavajicim resenim vyhledavani na konkrtnim priklade Diskuse mozneho budouciho rozsirovani

4 Závěr

V této diplomové práci jsem nejprve našel zdroje dat...

4.1 Dosažení vytčených cílů

...

Literatura

- [1] STROSSA, Petr. *Počítačové zpracování přirozeného jazyka*.
Vyd. 1. Praha: Oeconomica, 2011 000 s. ISBN 978-80-245-1777-3.
- [2] AYSE Göker a DAVIES John. *Information Retrieval: Searching in the 21st Century*.
Library of Congress Cataloging-in-Publication Data, 2009. ISBN: 978-0-470-02762-2.
- [3] AGGARWAL Charu C., ZHAI ChengXiang. *Mining Text Data*.
Springer New York Dordrecht Heidelberg London, 2000. ISBN 978-1-4614-3222-7.
- [4] CLINTON Gormley, ZACHARY Tong. *Elasticsearch: The Definitive Guide*.
O'Reilly Media, 2015. ISBN 978-1-4493-5854-9.
- [5] KENNEDY William. *Go in Action*.
Manning Publications Co, 2016. ISBN 978-1-6172-9178-4.
- [6] Elasticsearch. *Elasticsearch Reference* [online].
2017 [cit. 2017-03-03]. Dostupné z:
<https://www.elastic.co/guide/en/elasticsearch/reference/current/index.html>

Rejstřík

...

A Obsah přiloženého DVD

- Soubor Diplomova_prace_2017_Ludek_Vesely.pdf
 - Text diplomové práce