

**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

<b>1</b>	<b>Úvod</b>	<b>8</b>
1.1	Cíle práce . . . . .	8
1.2	Cílová skupina . . . . .	8
1.3	Použité metody . . . . .	8
1.4	Struktura práce . . . . .	8
<b>2</b>	<b>Komentovaná rešerše informačních zdrojů</b>	<b>9</b>
<b>3</b>	<b>Analýza byznys požadavků</b>	<b>10</b>
<b>4</b>	<b>Teorie fulltextového vyhledávání</b>	<b>11</b>
<b>5</b>	<b>Návrh řešení</b>	<b>12</b>
5.1	Nastroje použitelné pro fulltextové vyhledávání . . . . .	12
5.2	Návrh aplikace . . . . .	12
<b>6</b>	<b>Implementace</b>	<b>13</b>
6.1	Architektura aplikace . . . . .	13
6.2	Nastavení a nasazení Elasticsearch . . . . .	13
6.3	Backend . . . . .	13
6.4	Frontend . . . . .	13
6.5	Deployment . . . . .	13
6.6	Testování, overení . . . . .	13
<b>7</b>	<b>Závěr</b>	<b>14</b>
7.1	Dosažení vytčených dílů . . . . .	14
<b>A</b>	<b>Obsah přiloženého DVD</b>	<b>17</b>

# **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 Analýza byznys požadavků na aplikaci**

- Vysvetlení problému, který je resen
- 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...)

## **4 Teorie fulltextového vyhledávání**

- Popis problému indexování dokumentu
- Analýza dat použitých k vyhledávání
- Zpracování přirozeného jazyka
- Sklonování, časování, tvaroslovi obecně
  - Vaha slov, stop slova
  - Preklepy
  - Synonyma, zargon
  - Lemmatizace
  - Tezaurus

## **5 Návrh řešení**

### **5.1 Nastroje použitelné pro fulltextové vyhledávání**

- Popis a porovnání vhodných nástrojů
  - MySQL
  - Elasticsearch
  - Sphinx
  - Algolia
- Návrh architektury aplikace

### **5.2 Návrh aplikace**

- Definice případu užití
- Návrh architektury aplikace

## **6 Implementace**

- Vyber technologii a nastroju (jazyk, framework...)

### **6.1 Architektura aplikace**

- Frontend vs backend (API)

### **6.2 Nastavení a nasazení Elasticsearch**

- Implementace indexace
- Implementace vyhledávání

### **6.3 Backend**

- Popis tříd, rozhraní, implementační detaily

### **6.4 Frontend**

- Frontend: UI, komunikace s backendem

### **6.5 Deployment**

- Deployment, CI

### **6.6 Testování, overení**

- Overení funkčnosti a přínosu
- Testování, zda jsou splněna akceptační kritéria
- Porovnání s stavajícím řešením vyhledávání na konkrétním příkladě

## **7 Závěr**

V této diplomové práci jsem...

Diskuse možného budoucího rozšiřování

### **7.1 Dosažení vytčených cílů**

...

# Literatura

- [1] STROSSA, Petr. *Počítačové zpracování přirozeného jazyka*.  
Vyd. 1. Praha: Oeconomica, 2011 316 s. ISBN 978-80-245-1777-3.
- [2] AYSE Göker a DAVIES John. *Information Retrieval: Searching in the 21st Century*.  
Vyd. 1. Library of Congress Cataloging-in-Publication Data, 2009 295 s.  
ISBN: 978-0-470-02762-2.
- [3] AGGARWAL Charu C., ZHAI ChengXiang. *Mining Text Data*.  
Springer New York Dordrecht Heidelberg London, 2000 522 s. ISBN 978-1-4614-3222-7.
- [4] CLINTON Gormley, ZACHARY Tong. *Elasticsearch: The Definitive Guide*.  
O'Reilly Media, 2015 724 s. ISBN 978-1-4493-5854-9.
- [5] KENNEDY William. *Go in Action*.  
Manning Publications Co, 2016 264 s. 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