



**ČESKÉ VYSOKÉ  
UČENÍ TECHNICKÉ  
V PRAZE**

**F3**

**Fakulta elektrotechnická  
Katedra počítačů**

**Semestrální projekt**

# **Sémantické facetové vyhledávání na platformě React**

**Daniel Bourek**

**Softwarové inženýrství a technologie**

**Leden 2022**



## Poděkování / Prohlášení

poděkování

Prohlašuji, že jsem předloženou práci  
vypracoval samostatně.

# Abstrakt / Abstract

abstrakt

abstract

## / Obsah

<b>1 Úvod</b>	<b>1</b>
<b>2 Úvod do sémantických technologií</b>	<b>2</b>
2.1 Historie, Web verze, . . . . .	2
2.2 Vývoj v posledních letech . . . . .	2
<b>3 Facetové vyhledávání</b>	<b>3</b>
3.1 Popis . . . . .	3
3.2 Typy facetů . . . . .	4
3.2.1 Basic facet . . . . .	4
3.2.2 Range facet . . . . .	4
3.2.3 Bucket facet . . . . .	4
3.3 Srovnání přístupů . . . . .	4
3.3.1 Staticky definované faceti . . . . .	4
3.4 Převod do sémantického světa . . . . .	4
3.5 Osvědčené postupy . . . . .	4
<b>4 Popis implementace</b>	<b>5</b>
4.1 <a href="https://www.npmjs.com/package/@tpluscode/sparql-builder">https://www.npmjs.com/package/@tpluscode/sparql-builder</a> . . . . .	5
<b>Literatura</b>	<b>6</b>



# Kapitola 1

## Úvod

Málokterý vynález ovlivnil svět v takové míře jako vznik World Wide Web (zkráceně WWW či web). Za poměrně krátkou dobu své existence se web rozšířil téměř do každé části našeho života a dnes si bez něj lze svět jen těžko představit. Oproti ostatním ICT technologiím, které se často výrazně inovují a mění každých několik let, web funguje už 20 let téměř stejně. To se však začíná měnit s příchodem Sémantického webu, který zásadně ovlivňuje, jak přistupujeme k datům v internetu – místo relací mezi dokumenty (hypertextové odkazy) můžeme vytvářet relace mezi fakty. Svět lze tak mnohem lépe popsat a stává se pro nás srozumitelnější. Navíc jsou tyto relace více strojově čitelné, tudíž můžeme nad těmito daty mnohem přesněji vyhledávat informace či vykonávat automatizace.

Tato změna si žádá nové přístupy k ukládání, zpracování a vyhledávání dat. Právě vyhledáváním v sémantických datech se bude zabývat tato práce, konkrétně facetovým vyhledáváním. Facetové vyhledávání, tedy zatřídění vyhledaných výsledků do různých kategorií, je v současné době velmi rozšířené. Pomáhá nám upřesnit výsledky vyhledávání a najdeme jej tedy skoro v každém větším e-shopu. Přístupů k facetovému vyhledávání je více, ne všechny jsou však vhodné pro sémantická data. Nad sémantickými daty tak existuje velmi málo řešení facetového vyhledávání a k tomu jsou často závislé na nějaké platformě. Cílem této práce je tak:

1. Srovnat existující přístupy k facetovému vyhledávání, především pak z hlediska využití sémantických technologií.
2. Navrhnout modul sémantického facetového vyhledávače, který bude umožňovat rozdělení vyhledávání a jeho vizualizace do samostatných modulů.
3. Naimplementovat prototyp modulu sémantického vyhledávače.

# Kapitola 2

## Úvod do sémantických technologií

### 2.1 Historie, Web verze,

podobný vysvětlení jako ve videu [??]

### 2.2 Vývoj v posledních letech

tady dát jak je to jen 10 let starý vlastně  
velký pokrok v RDF, OWL a SPARQL  
odkazy k částem dokumentů + google vyhledávání taky a



# Kapitola 3

## Facetové vyhledávání

V této kapitole si popíšeme co je facetové vyhledávání a k čemu se primárně využívá. Zanalyzujeme a srovnáme pak různé přístupy k implementaci facetového vyhledávání, především z hlediska využití sémantických technologií. Abychom získali přehled o používaných řešeních facetového vyhledávání, zanalyzujeme poskytované Facet Search APIs největších společností v této oblasti jako Elastic či Solr.

### 3.1 Popis



Obrázek 3.1. Ukázka facetů s vysvětlivkami. [1]

Facetové vyhledávání je zatřídění vyhledaných výsledků do různých kategorií (facetů) dle kterých se dá sada výsledků dále filtrovat. Dá se ním tak obohatit každé vyhledávání, ale často bývá spojeno s fulltextovým vyhledáváním (TODO možná někde vysvětlit co to je), aby uživateli umožnilo jeho dotaz dále upřesnit. Hojně se využívá třeba v e-commerce sektoru, kde podle studie Nielsen Norman Group (NNG) z roku 2018 jsou e-shopy bez facetového vyhledávání výjimkou.[2] Jelikož není definovaný žádný standard facetového vyhledávání, zadefinujeme si, co by měl takový modul facetového vyhledávání splňovat:

- facet obsahuje hodnotu pro každý výsledek ze sady výsledků
- jednotlivé facety lze kombinovat mezi sebou
- mezi kritérii facetů platí logický AND (ne pouze OR), tzn. aby se výsledek objevil v sadě výsledku, musí vyhovět všem aktivním facetům
- hodnoty facetů ukazují počet výsledků, které aplikování facetu s danou hodnotou v aktuálním stavu vrátí

- hodnoty facetů, které by vrátily prázdnou sadu výsledků se nezobrazují nebo jsou „disabled“<sup>1</sup>

## 3.2 Typy facetů

Typy facetů řadíme dle toho jaké hodnoty a jakým způsobem jdou navolit.

### 3.2.1 Basic facet

Základní typ facetu s možností volby nejvýše jedné hodnoty podle které je pak sada výsledků filtrována. Ovládacím prvkem bývá select element.

### 3.2.2 Checkbox facet

Facet s možností volby více hodnot skrz zaškrtování checkboxů.

### 3.2.3 Range facet

Facet pro číselná data s možností nastavení rozsahu. Ovládacím prvkem bývá posuvník (input element s hodnotou atributu type range).

### 3.2.4 Bucket facet

Podobné jako range facet, akorát se neovládá posuvníkem, ale jsou nadefinovány rozsahy, do kterých se pak výsledky roztřídí.

## 3.3 Srovnání přístupů

Řešení jak implementovat facetové vyhledávání je více. Obecně je lze rozdělit podle toho, kde dochází k filtraci výsledků aktivními facety, tedy buďto na straně klienta či na straně serveru. U implementace facetového vyhledávání je také nutné myslet na to jakým způsobem se budou plnit data facetů. Nejčastěji se setkáváme s tím, že se hodnoty facetů posílají ve stejné response jako sada výsledků. TODO modul vs server ještě

### 3.3.1 Filtrování na straně klienta

### 3.3.2 Filtrování na straně serveru

search results obsahují facety eg. <https://www.npmjs.com/package/@ebi-gene-expression-group/scxa-faceted-search-results>

## 3.4 Převedení do sémantického světa

.

<sup>1</sup> Jejich HTML ovládací prvek má atribut disabled.



## **Kapitola 4**

### **Popis implementace**



#### **4.1**

**<https://www.npmjs.com/package/@tpluscode/sparql-builder>**



## Literatura

- [1] *Semantic University*. 2021.  
<https://cambridgesemantics.com/blog/semantic-university/intro-semantic-web>.
- [2] *What is faceted search*. 2018.  
<https://stackoverflow.com/questions/5321595/what-is-faceted-search>.
- [3] *The State of Ecommerce Search*. 2018.  
<https://www.nngroup.com/articles/state-ecommerce-search>.