



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

GENERÁTOR VĚDECKÝCH WEBOVÝCH PORTÁLŮ

SCIENTIFIC WEB PORTAL GENERATOR

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

JIŘÍ FURDA

VEDOUCÍ PRÁCE

SUPERVISOR

doc. RNDr. PAVEL SMRŽ, Ph.D.

BRNO 2019

Abstrakt

Do tohoto odstavce bude zapsán výtah (abstrakt) práce v českém (slovenském) jazyce.

Abstract

Do tohoto odstavce bude zapsán výtah (abstrakt) práce v anglickém jazyce.

Klíčová slova

Python, Elasticsearch, Vue, Flask, fasetové vyhledávání

Keywords

Python, Elasticsearch, Vue, Flask, faceted search

Citace

FURDA, Jiří. *Generátor vědeckých webových portálů*. Brno, 2019. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce doc. RNDr. Pavel Smrž, Ph.D.

Generátor vědeckých webových portálů

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana doc. RDNr. Pavla Smrže, Ph. D. Další informace mi poskytl... Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Jiří Furda
1. května 2019

Poděkování

V této sekci je možno uvést poděkování vedoucímu práce a těm, kteří poskytli odbornou pomoc (externí zadavatel, konzultant, apod.).

Obsah

1	Úvod	3
2	Rozbor řešené problematiky	4
2.1	Webové portály	4
2.1.1	CORDIS	4
2.1.2	Funding and Tenders	5
2.2	Vývoj webu	5
2.2.1	MVC architektura	5
2.2.2	Responzivita	6
2.3	Bootstrap	7
2.4	Python	8
2.4.1	Flask	8
2.5	Vue.js	8
2.5.1	Vue Router	9
2.5.2	Nuxt.js	9
2.5.3	Vuex	9
2.6	Elasticsearch	11
2.6.1	Základní struktura	11
2.6.2	Dotazy	12
2.6.3	Metadata	12
2.6.4	Fasetové vyhledávání	13
2.6.5	Plnotextové vyhledávání	13
2.6.6	Podobnostní hledání	14
2.6.7	Knihovny	15
3	Návrh a implementace systému	16
3.1	Extrakce dat	16
3.2	Webový portál	16
3.2.1	Návrh	17
3.2.2	Schéma	17
3.2.3	Podoba	18
3.2.4	Struktura adresáře	19
3.2.5	Vue komponenty	19
3.2.6	Vyhledávání	20
3.2.7	Vyhledávání hodnot faset	22
3.2.8	Model fasety	23
4	Experimenty a vyhodnocení	24

4.1	Testování na uživateli	24
4.2	Zátěžové testy	24
4.3	Rychlost obnovy dat	25
4.4	Možnost úprav	25
5	Závěr	27
5.1	Možnosti pokračování	27
	Literatura	28

Kapitola 1

Úvod

Webové portály Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Pro toto téma jsem se rozhodl, protože mám k webovým technologiím blízko již od dětství a chci v tomto směru nadále rozvíjet své znalosti.

Cílem je Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Tato práce vychází z bakalářské práce *Automaticky aktualizovaný webový portál* Petra Staňka [5]. Webový portál byl od základu implementován znovu s použitím modernějších technologií. Extrakce dat zůstala zachována, pouze byla rozšířena o extrakci témat výzev k předkládání návrhů.

V první kapitole Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Kapitola 2

Rozbor řešené problematiky

2.1 Webové portály

Webový portál je druh webové stránky, která shromažďuje informace z vícero různých zdrojů a uživateli ty nejrelevantnější informace prezentuje uživateli shromážděné na jednom místě [3].

Zpravidla je umožněno na portálu v těchto informacích vyhledávat. Velmi časté je také zakomponování autentizace uživatele a podle jeho role jsou mu zpřístupněny různé části daného portálu [7].

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.1.1 CORDIS

Tato kapitola čerpá z [2]. CORDIS (*The Community Research and Development Information Service*) je portál provozovaný Evropskou Unií (dále jen EU) sloužící jako hlavní zdroj výsledků projektů, sponzorovaných v rámci programů EU pro výzkum a inovaci. Na jednom místě veřejně poskytuje informace jak o těchto projektech, tak i o jejich účastnících, o hlášeních, vědeckých zprávách a publikacích. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin.

Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.1.2 Funding and Tenders

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.2 Vývoj webu

Tato kapitola čerpá z [1].

Tvorba webových stránek se obecně rozlišuje do dvou celků zvaných *frontend* a *backend*.

Frontendová část webu je to, co uživatel vidí, když stránku navštíví. Sestavení této části se v žádném případě neobejde bez značkovacího jazyka HTML (*Hypertext Markup Language*), který udává strukturu a obsah webové stránky. Další nedílnou součástí jsou kaskádové styly neboli CSS (*Cascading Style Sheets*). Ty se starají o to, jak bude obsah stránky esteticky působit. Další součástí, se kterou se setkáte téměř na každém webu je skriptovací jazyk JavaScript, který umožňuje webové stránky obohatit o interaktivitu. Zpracování této části webu probíhá v klientovi - tedy internetovému prohlížeči uživatele. Na základě toto faktu se lze každé stránce na internetu „nahlédnout pod pokličku“ a zdrojové kódy frontendové části zobrazit, ve většině případů i rovnou v internetovém prohlížeči upravit, což má samozřejmě vliv jen pro současného uživatele a ostatním návštěvníkům se tyto změny neprojeví.

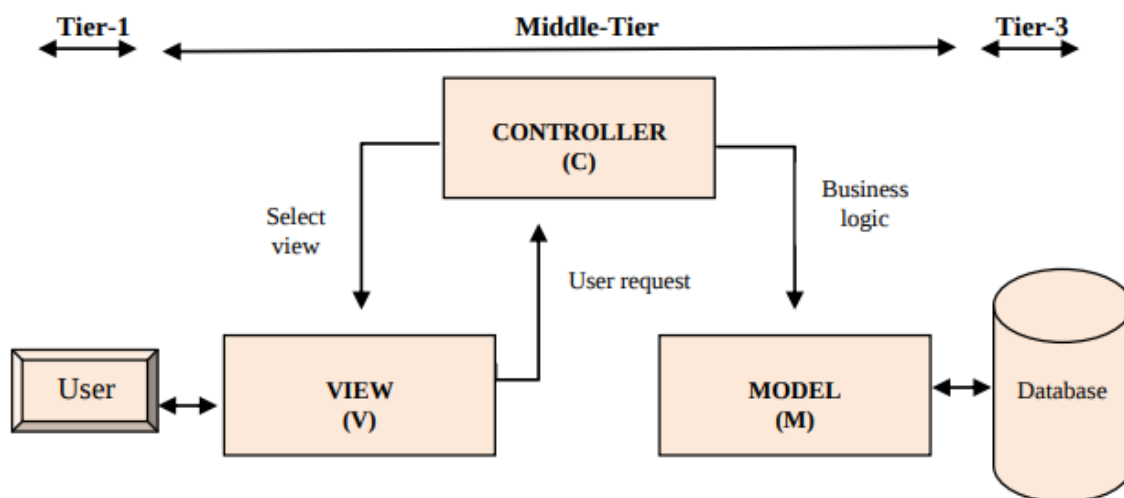
Oproti tomu backendová část je od uživatele izolována a zdrojové kódy pro něj nejsou přístupné, pokud je sám autor někde nezveřejní. Jejich interpretace totiž probíhá přímo na serveru a ten poté klientovi odešle již zpracovanou frontendovou část. Výběr technologie implementace backendové části je podstatně rozmanitější, patří mezi ně například Python, PHP, Java, C, C++, .NET, Ruby a další. Díky prostředí Node.js¹ lze ale využít i dříve zmiňovaný JavaScript, který sám o sobě pro tento účel není primárně určen.

2.2.1 MVC architektura

Většinu aplikací lze obecně rozdělit na tři hlavní celky - data, rozhraní a logiku. V anglickém jazyce se tyto celky dají nazvat pomocí slov *model*, *view* a *controller* (zkratka MVC).

¹Node.js: <https://nodejs.org/en/>

Historie této architektury sahá až do sedmdesátých let, nicméně je v dnešním světě stále naprostým standardem vývoje.



Obrázek 2.1: Diagram toku akcí při využití MVC architektury

- Pojmem pohled (*view*) se označuje uživatelské rozhraní (například stránka zobrazená v prohlížeči), tato vrstva prezentuje data a je to také jediná vrstva, s kterou uživatel přímo komunikuje. Akce provedené v pohledu se předávají kontroleru ke zpracování.
- Kontroler (*controller*) se chová jako prostředník mezi daty a rozhraním, zpracovává požadavky uživatele a provede s modelem potřebné akce. Ve většině případu po provedené akci kontroler znovu obnoví pohled.
- Model je obvykle objekt obsahující patřičná data z databáze. Zprostředkovává jak jejich získání, tak i ukládání. Mimo to může být objekt obohacen i o funkce, zpracovávající tyto data (například může převádět Unixový čas do formátu lépe čitelného pro člověka).

Oddělením těchto celků je zajištěna lepší organizovanost zdrojových kódů a tím je usnadněn rychlejší vývoj. Jednotlivé celky jsou také lehčeji znovupoužitelné a práce programátorů na nich může probíhat souběžně. Rozdělením je také možné pro jeden pohled mít pohledů hned několik. Další pohledy mohou přibývat nebo naopak ty stávající mohou být smazány a na databázi aplikace to nebude mít žádný vliv.

2.2.2 Responzivita

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit

amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.3 Bootstrap

Bootstrap² je frontendový framework pro snadný vývoj webových stránek. Je vyvíjen společností Twitter, která jej v roce 2010 vydala jako otevřený software. Jen dva roky poté se Bootstrap stal nejoblíbenějším projektem na stránce GitHub. Bootstrap se skládá převážně z kaskádových stylů, ale interaktivní prvky využívají programovací jazyk JavaScript s knihovnou jQuery. Pro interaktivní část ale existují i další alternativy využívající například Vue.js³, React⁴ nebo Angular⁵.

Bootstrap se drží moderních zásad jako například přístup mobile-first.

- Drží se přístupu mobile-first
- Nabízí podporu mnoha prohlížečů
- Umožňuje tvořit responzivní design

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

²Bootstrap: <https://getbootstrap.com/>

³BootstrapVue: <https://bootstrap-vue.js.org/>

⁴React Bootstrap: <https://react-bootstrap.github.io/>

⁵NG Bootstrap: <https://ng-bootstrap.github.io/#/home/>

2.4 Python

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.4.1 Flask

Flask je open source webový micro-framework napsaný v programovacím jazyce Python. Klíčový znak micro-frameworků je co nejmenší nebo i nulová závislost na externích knihovnách. Flask si klade za cíl udržet co nejjednodušší jádro, ale zároveň se soustředí na možnost tento základ jednoduše rozšířit. Smyslem Flasku je být základním kamenem pro jakoukoliv webovou aplikaci. Z tohoto důvodu na rozdíl od ostatních frameworků neobsahuje žádnou abstrakci databázové vrstvy nebo knihovnu pro formuláře. Tohoto lze docílit použitím některého z rozšíření.

Ve svém základu využívá šablonovací jazyk Jinja2⁶ zjednodušující udržení konzistentní struktury stránek. Poskytuje například dědičnost šablon nebo znovupoužitelné bloky a dokáže se vypořádat s bezpečnostními útoky typu XSS (Cross Site Scripting). Druhou knihovnou, kterou Flask využívá je Werkzeug - jedna z nejpokročilejších knihoven pro rozhraní brány webového serveru (anglicky WSGI neboli Web Server Gateway Interface), které umožňuje webovou aplikaci provozovat nezávisle na technologii použité na serveru.

Tento framework využívá například sociální síť Pinterest⁷ [6] nebo síťová dopravní společnost Lyft⁸ [4] (podobná službě Uber⁹).

2.5 Vue.js

V dnešní době jsou webové prohlížeče stále schopnější a výkonnější. Díky tomu se stává trendem přenášet stále větší části webové aplikace ze serveru na stranu klienta. Toho je docíleno pomocí programovacího jazyku JavaScript. V současné době se nejvíce využívá

⁶Jinja2: <http://jinja.pocoo.org/>

⁷Pinterest: <https://cz.pinterest.com/>

⁸Lyft: <https://www.lyft.com/>

⁹Uber: <https://www.uber.com/cz/cs/>

knihovna jQuery, ale začali se objevovat pokročilejší frameworky jako například React, Angular a v neposlední řadě Vue.

Vue je progresivní framework - přizpůsobuje se složitosti projektu. Podobně jako dříve zmiňovaný Flask se nejedná o rozsáhlý framework, jehož části by jste vypínali, ale jedná se o základní framework, který můžete dále rozšiřovat. Díky tomuto přístupu učební křivka není tak strmá, jako u konkurenčních frameworků. Jediné, co člověku pro začátek práce s Vue stačí je být obeznámen s HTML a s čistým JavaScriptem.

Jedna z předností toto frameworku je jednoduchost obousměrné vazby dat. Tato vazba znamená, že hodnota proměnné v jazyce JavaScript je synchronizována s hodnotou v objektovém modelu dokumentu (DOM) a to stejné platí i v opačném směru. V praxi může být tato funkce využita například v internetovém obchodě, kdy uživatel klikne na tlačítko „Přidat do košíku“, které pouze rozšíří pole košíku o další produkt. Bez jakéhokoliv dalšího úsilí komponenty, závislé na této proměnné zaznamenají změnu a automaticky provedou odpovídající akce - košík přepočítá celkový počet vybraného zboží, celkovou cenu nákupu, zlevní se cena daného produktu, při koupi dalších kusů apod.

2.5.1 Vue Router

Mezi oficiálně podporované knihovny patří například Vue Router, umožňující tvorbu jednostránkové aplikace, kdy se načte pouze jediná stránka a pomocí JavaScriptu a asynchroniích požadavků se mění části obsahu webu. Server v takovém případě slouží pouze jako úložiště dat a veškerou prezentaci se stará strana klienta.

2.5.2 Nuxt.js

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.5.3 Vuex

Dalším oficiálním rozšířením je Vuex, knihovna pro správu stavu. Pokud se v aplikaci používá více komponent, které využívají stejnou proměnnou, brzy se zdrojový kód stává neudržitelným. V tuto chvíli je vhodné využít knihovnu Vuex. Jejím účelem je udržovat centrální stav proměnných, které jsou sdíleny napříč různými komponentami v aplikaci. Tohoto je docíleno díky dodržování návrhového vzoru jménem Flux. Mezi základní pilíře návrhového vzoru Flux patří několik následujících pravidel, kterých se Vuex drží.

Jediný zdroj pravdy

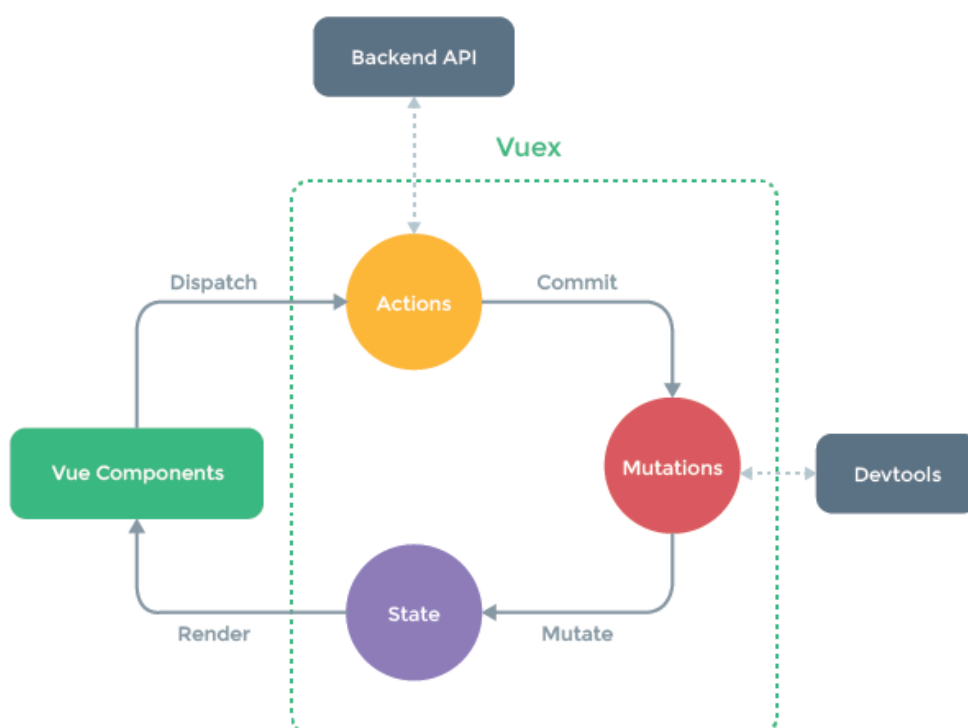
Data, která jsou sdílena více komponentami jsou uložena na jednom místě - sklad (anglicky *store*) a jsou oddělena od komponent, které je využívají. Komponenty mohou stále mít svá lokální data, ale nesmějí mít svou kopii sdílených dat, ty se vždy musí číst ze skladu.

Data pouze ke čtení

Komponenty nesmějí přímo upravovat data ve skladu. V případě potřeby změny těchto dat sklad pouze informují a ten se sám o postará o provedení změny pomocí tzv. mutací (Reprezentovány uzlem *Mutations* v grafu 2.2). Díky tomu se minimalizuje šance nepředpokládaných změn těchto dat a funkce, starající se o tyto změny jsou jednodušeji dohledatelné.

Změny dat jsou synchronní

Asynchronní provádění operací mnohdy přináší spoustu výhod, v tomto případě je ale vyžadováno mutace provádět synchronně kvůli odstranění závislosti na pořadí a načasování různých událostí.



Obrázek 2.2: Diagram znázorňující životní cyklus dat v knihovně Vuex

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris.

Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

BootstrapVue

2.6 Elasticsearch

Elasticsearch je volně šiřitelný vyhledávač umožňující distribuované vyhledávání a analýzu v reálném čase. Umožňuje rychle pracovat s rozsáhlými daty (anglicky *big data*) a provádět nad nimi plnotextové vyhledávání, strukturované vyhledávání a nebo zmiňovanou analýzu.

Tento vyhledávač je hojně využíván jak menšími projekty, tak i světoznámými organizacemi. Za zmínku stojí například internetová encyklopedie Wikipedia¹⁰, britský deník The Guardian¹¹, komunitní stránka pro výměnu rad programátorů StackOverflow¹² nebo populární webová služba pro verzovací nástroj Git a to GitHub¹³.

Elasticsearch je postavený na knihovně Apache Lucene¹⁴, využívá především její funkce plnotextového vyhledávání, ale oproti této knihovně samotné Elasticsearch nabízí o mnoho přívětivější způsob používání, především pro začínající uživatele.

Jedná se o bezschémovou databázi, oproti relačním databázím nenabízí propojovací dotazy, a proto je nutné ukládaná data denormalizovat. Díky tomu jsou ale data a jejich metadata v těsné blízkosti a tím je docíleno rychlé fulltextové vyhledávání.

2.6.1 Základní struktura

Základní struktura databáze Elasticsearch se na první pohled od tradičních relačních databází velmi neliší, ke spoustě pojmů se dá najít ekvivalent.

Uzel a shluk

Pod pojmem uzel (anglicky *node*) se skrývá server, ve kterém jsou uložena data. Uzel v rámci databáze může být jeden nebo jich může být i více a každý bude obsahovat pouze určitou část dat. Tyto uzly jsou seskupeny do shluku (anglicky *cluster*).

Dokument

Dokument je základní jednotka dat, které může být zaindexována. Obsah dokumentu je zapsán ve formátu JSON. V relační databázi lze dokument přirovnat k řádku tabulky.

Index

Skupina dokumentů s podobnou strukturou se nazývá index. Ekvivalentem v relačních databázích by v tomto případě byla tabulka.

Střepy

Elasticsearch nabízí možnost index rozdělit na střepy (anglicky *shards*), které mohou být uloženy v různých uzlech. Tuto možnost je vhodné využít především při práci s indexy s

¹⁰Wikipedia: <https://www.wikipedia.org/>

¹¹The Guardian: <https://www.theguardian.com/international>

¹²StackOverflow: <https://stackoverflow.com/>

¹³GitHub: <https://github.com/>

¹⁴Apache Lucene: <https://lucene.apache.org/>

velkým obsahem dat. Díky rozdělení docílíme distribuování zátěže a lze využít paralelní zpracování, tudíž bude zpracování dotazu odbaveno výrazně rychleji. Přístup k takto rozděleným indexům se nijak nemění a pro uživatele je tento proces naprosto transparentní.

Repliky

Obdobně lze využít i repliky. Zde se ovšem data indexu nerozdělují, ale naopak duplikují. Díky uložení těchto replik mezi více uzly může systém být robustnější a dokáže fungovat i v případě selhání části shluku. I v tomto případě je umožněno vyhledávat paralelně.

2.6.2 Dotazy

Pro komunikaci s touto databází se využívá Query DSL - dotazy aplikačního rozhraní REST ve formátu JSON.

Obsah dotazu může vypadat například následovně

```
{
  "query": {
    "match": {
      "address": "todo"
    }
  }
}
```

Na tento dotaz můžeme od serveru dostat odpověď s obsahem

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Kontext

Dotazy v Elasticsearch se rozlišují na dva základní typy - kontext dotazu (*query*) a kontext filtru. První z nich se dá lidsky přeložit do otázky „Jak moc dokument splňuje dotaz?“. V tomto případě se u každého výsledku počítá skóre relevance, podle kterého jsou ve výsledné odpovědi dokumenty sestupně řazeny.

2.6.3 Metadata

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue

a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.6.4 Fasetové vyhledávání

S velkým množstvím dat přichází potřeba obsah podle určitých kritérií zúžit. K tomuto účelu slouží filtry a fasetové vyhledávání (někdy také označována jako fastová navigace - anglicky *faceted search* nebo *faceted navigation*). Využití faset je ale pokročilejší než použít filtry, protože fasetová navigace umožňuje využít hned několik filtrů najednou.

V Elasticsearch je možné tuto navigaci vytvořit pomocí tzv. kyblíkových agreagací (angl. *bucket aggregations*). Jednoduše řečeno, se pro dokumenty vytvoří kyblíky, kde každý z kyblíku odpovídá určitému kritériu. Pokud dokument toto kritérium splňuje, je do kyblíku vložen.

2.6.5 Plnotextové vyhledávání

Uložená data se dělí na dva typy - přesné hodnoty a plnotextové hodnoty (anglicky *full-text*). Přesné hodnoty jsou určeny pro data, jako například identifikátor uživatele. Pokud totiž vyhledáváme uživatele s identifikátorem 42, jako výsledek neočekáváme uživatele, který má identifikátor 420, ale očekáváme výsledek, který bude přesně splňovat hledanou položku. Pokud ale člověk pracuje s textovými položkami, nemusí nutně hledat stoprocentní shodu. Lidský jazyk je velice rozmanitý a umožňuje slova časovat, skloňovat, či vyjádřit jednu věc vícero synonyma. V těchto případech hledání totožných hodnot selhává, i když má text z pohledu člověka stejný nebo podobný význam. Pro tento typ hledání slouží právě plnotextové hledání. Můžeme se s ním setkat například při hledání článku v internetovém magazínu. Elasticsearch pro takové hledání musí text nejprve analyzovat a vytvořit pro něj obrácené indexy.

Obrácený index

Obrácený index (anglicky *inverted index*) je struktura obsahující seznam všech unikátních slov vyskytujících se v uložených dokumentech. Pro každé z těchto slov uchovává informaci, ve kterých dokumentech se vyskytovalo. Díky tomu je možné v textových položkách rychle provádět plnotextové hledání.

Příkladem mohou být dva dokumenty, kde každý obsahuje právě jednu z následujících vět.

-
-

Zpracování těchto dokumentů tedy může vzniknout následující obrácený index. Takovýto index by byl ale v praxi nepoužitelný. Text nestačí pouze strojově rozdělit do tokenů, ale musí proběhnout celá analýza textu.

Analýza

Analýza představuje proces, kdy se kus textu rozdělí do jednotlivých tokenů a následně se tyto tokeny upraví do podoby vhodné pro použití plnotextového vyhledávání. Tento proces obstarává analyzátor, který se skládá ze tří částí:

- *Filtry znaků* upraví zpracovávaný řetězec ještě před samotným rozdělením do tokenů. Jeho úkolem může například být odstranění HTML značek z textu.
- *Tokenizér* z textu sestaví jednotlivé tokeny. Tokeny ve většině případů představují jednotlivá slova, ale nemusí tomu tak vždy být.
- *Filtry tokenů* na konec projde všechny vytvořené tokeny, může některé odstranit, přidat a nebo jen upravit.

Filtrů může být v jednom analyzáru použitá celá řada a to jak filtrů znaků, tak i filtrů tokenů. Tokenizér je ale v rámci analyzáru vždy použit jen jeden jediný.

Elasticsearch již ve svém základu nabízí škálu různých analyzáru. Tyto analyzéry jsou různé složitosti, mezi ty nejjednodušší patří analyzér bílých znaků, který text rozdělí striktně podle mezer, tudíž je i tečka na konci věty vnímána jako část slova. Nejčastěji se ale člověk může setkat s jazykovým analyzérem, který je uzpůsoben lidské řeči. Elasticsearch takový analyzér nabízí pro mnoho jazyků mezi které patří i jazyk český.

Filtry tokenů

Velmi častou filtrací bývá změna velkých písmen na malé, díky tomu nejsou slova na začátku věty vnímána jako jiná slova, než ty v jiné části věty. Pro český jazyk je také typické odstranění diakritiky.

Součástí analyzátorů bývá často odstranění tokenů, které obsahují nepodstatná slova. Jedná se především o spojky a předložky. Taková slova se v plnotextovém vyhledávání označují jako *stop slova* (anglicky *stop words*). Tento seznam je vždy specifický pro konkrétní jazyk textu, použít výchozí anglická stop slova nad česky psaným textem by ztrácelo význam.

Další nedílnou součástí pokročilejších analyzátorů je *stematizace*, aneb převod slov do jejich základního tvaru. Pro tuto operaci lze využít dva způsoby a to algoritmus nebo slovník. Každý z těchto způsobů má své pro a proti. Využitím algoritmické stematizace sice analyzátor nemusí znát všechna slova v daném jazyce, protože slova převádí do základního tvaru pouze na základě sady pravidel, ale zase dochází k určité nepřesnosti. Elasticsearch v případě využití české algoritmické stematizace ze slov odstraňuje všechny přípony. Druhým způsobem, který lze využít je stematizace na základě slovníku. Díky tomu jsou převedené tvary přesnější než ty získané algoritmem, ale hlavním předpokladem je využití vhodného slovníku s dostatečnou zásobou slov.

2.6.6 Podobnostní hledání

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo,

lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.6.7 Knihovny

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Kapitola 3

Návrh a implementace systému

Předchozí řešení mých kolegů umožňovalo vyhledávat pouze projekty a vědecké zprávy. Nabízelo se tento obsah rozšířit o výzvy k předkládání návrhů, na základě kterých Evropská komise vybírá organizace či fyzické osoby pro uskutečnění projektů se spolufinancováním Evropské unie. Samotné výzvy jsou ale hodně obecné a zaštiťují mnoho témat napříč různými oblastmi. Dalším úskalím byl nedostatek informací v popisu výzev. Rozsáhlý slovní popis se vztahovat až ke konkrétním tématům. Proto jsem se rozhodl portál nerozšiřovat o výzvy ale právě o zmiňovaná témata.

3.1 Extrakce dat

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Pro tyto účely jsem využil aplikační rozhraní nabízené přímo portálem Fundings and Tenders Rozhraní nabízí potřebná data ve formátu JSON, čímž bylo značně zjednodušeno jejich dolování.

3.2 Webový portál

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

3.2.1 Návrh

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

3.2.2 Schéma



Obrázek 3.1: TODO

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor

sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

3.2.3 Podoba

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.



Obrázek 3.2: Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullam-

corper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

3.2.4 Struktura adresáře

Veškeré zdrojové kódy portálu se nachází v adresáři **portal**. Uvnitř této složky se nachází adresáře odpovídající architektuře MVC.

- Složka **models** obsahuje modely
- Složka **controllers** obsahuje kontrolery
- Složka **templates** obsahuje pohledy
- Složka **static** obsahuje kaskádové styly a zdrojové kódy Vue komponent
- Soubor **app.py** sloužící ke spuštění portálu

3.2.5 Vue komponenty

Některé z využívaných faset obsahují i stovky položek a jejich hodnoty mají často dlouhé názvy. Jejich zobrazení pouze v postranním panelu jsem považoval za nedostatečné, proto jsem se rozhodl toto řešení oproti předchozímu řešení portálu přepracovat s využitím technologie Vue.js (představenou v kapitole 2.5) a několika prvků z frameworku BootstrapVue¹. Pro každou z faset jsem vytvořil dvě komponenty - položku v postranním panelu a modální okno.

Postranní panel

Tento panel je tvořen Vue komponentou **sidebar-facet-list**, která pro každou fasetu vytvoří instanci komponenty **sidebar-facet**.

Každá komponenta **sidebar-facet** ve výchozím stavu nabízí nejvýše pět nejčastěji vyskytujících se hodnot dané fasety. Ty jsou postupně nahrazovány zvolenými hodnotami dané fasety. Další hodnoty s menším výskytem jsou dostupné v modálním okně, které se zobrazí po kliknutí na tlačítko „More...“. Obsahem této komponenty je mimo jiné dílčí komponenta

¹BootstrapVue: <https://bootstrap-vue.js.org/>

frameworku BootstrapVue `b-collapse`, která v případě potřeby umožňuje danou fasetu minimalizovat. Pokud se takto stane a faseta má zvolené některé ze svých možností, objeví se vedle jejího názvu počet vybraných možností, aby měl uživatel i při minimalizovaných fasetách přehled o tom, které možnosti filtrování právě zvolil.



Obrázek 3.3: TODO

Modální okno

Okna jednotlivých faset jsou tvořena instancemi Vue komponenty `modal-facet`. O jejich vytvoření se stará komponenta `modal-facet-list`. Rozsáhlejší výpis možností daného fasetu se nyní zobrazují v modálním okně. Zde je možnost současně vidět podstatně více možností a také je možné vypsát jejich hodnoty v plném znění. V tomto okně je také obsaženo vstupní pole, které asynchronními dotazy na databázi Elasticsearch vyhledává mezi možnostmi daného fasetu. Samotné modální okno komponenta frameworku BootstrapVue, rozšířená o prvky nutné pro tento portál (například zmiňované vyhledávací pole)

Fasetová možnost

Jednotlivé fasetové možnosti jsou reprezentovány komponentou `option-facet`, která zobrazuje zaškrťovací pole, název dané možnosti a počet výsledků aktuálního vyhledávání obsahující tuto možnost. Při zaškrtnutí možnosti se její data zkopírují do proměnné `checkedOptions` příslušného fasetu. Tato komponenta se vyskytuje jak v modálním okně, tak i v postranním panelu.

Inicializace dat

Na každé stránce portálu se vyskytuje právě jednou komponenta `init-facet-data`, která jednorázově obstará naplnění proměnných programovacího jazyka JavaScript.

3.2.6 Vyhledávání

Samotné vyhledávání obstarává třída `IndexSearch`, při svém vytvoření uloží vstupní parametry a sestaví dotazy pro Elasticsearch. Mezi tyto parametry patří

- `index`, ve kterém se bude hledat
- pole dokumentu, ve kterém budou v případě shody zvýrazněna klíčová slova
- pole dokumentu, ve kterých se budou klíčová slova hledat

buildSearch()

První z dotazů je vytvořen touto metodou, která zpracuje všechny GET parametry protokolu HTML a v případě shody jejich názvů s některým z faset, přidá tuto hodnotu do dotazu. Tento dotaz se využívá pro hledání v modálním okně fasetu.

buildAggregationsSearch()

Druhý dotaz vzniká rozšířením dotazu z metody **buildSearch()**. Je obohacen o agregaci fasetových hodnot, které jsou poté zobrazeny v postranním panelu.

prepareLayoutData()

Poté co byly výše zmíněné dotazy vykonány je třeba výsledky zpracovat a připravit pro předání do odpovídajících Vue komponent. Výsledkem této funkce jsou tři struktury JSON, obsahující data využívané pro inicializaci komponent v šabloně, zabalení ve slovníku.

Jedna ze struktur, které je poté uložena do skladu knihovny Vuex může vypadat například následovně

```
{
  "facets": [
    {
      "mostFrequentOptions": [
        {
          "count": 8748,
          "text": "United Kingdom",
          "value": "United Kingdom"
        },
        {
          "count": 5276,
          "text": "Germany",
          "value": "Germany"
        },
        ...
      ],
      "field": "coordinator.country.keyword",
      "name": "coordcountry",
      "checkedOptions": [
        {
          "count": 4692,
          "text": "Spain",
          "value": "Spain"
        },
        ...
      ],
      "title": "Coord. Country"
    },
    ...
  ]
}
```


Každá z faset má svůj objekt obsažený v poli s klíčem **facets**.

- **mostFrequentOptions** obsahuje pole s nejvýše pěti nejčastějšími hodnotami fasety
- **name** značí unikátní interní název, díky kterému je portál schopný fasetu identifikovat
- **checkedOptions** je pole, kde se udržují zvolené hodnoty fasety
- **title** nese název fasety, který je zobrazován uživateli

Každá z hodnot faset je zase reprezentována objektem o třech attributech

- **count** udává množství výsledků hlavního hledání s výskytem této hodnoty fasety
- **text** je název hodnoty, která se zobrazuje uživateli
- **value** obsahuje reálnou hodnotu, která se aplikuje do dotazu hledání. Od atributu **text** se mnohdy neliší, ale je třeba mít tyto dvě informace rozdělené například pro použití možností s názvem sloučených ze dvou polí databáze (příkladem může být název a zkratka instituce)

createForIndex()

Protože se v portálu vytváří hledání na vícero místech, byla vytvořena statická metoda třídy **IndexSearch**, která vrací předpřipravenou instanci této třídy pro hledání v jednom z typů obsahu portálu. Tato funkce očekává ve svém jediném parametru řetězec, určující typ hledání (**projects**, **deliverables** nebo **topics**)

Hromadné akce

Pro případy, kdy je potřeba pracovat s několika hledáními naráz, si lze práci ušetřit využitím dalších statických metod třídy **IndexSearch** a to konkrétně metoda **createForEveryIndex()**, která vrací slovník se všemi třemi možnými typy hledání, a poté také **executeMany()**, která vykoná hledání všech instancí **IndexSearch** ve slovníku předaném této funkci v parametru.

3.2.7 Vyhledávání hodnot faset

Protože ve fasetách lze asynchronně vyhledávat, bylo nutné napsat aplikační rozhraní propojující databázi Elasticsearch s Vue komponentou reprezentující toto vyhledávání. Tuto činnost obstarává funkce **showApi** v kontroleru **facet_controller**.

Jako parametr, který je předán cestou URL adresy očekává platný název fasety, pro kterou existuje model třídy **Facet**. Další parametry jsou předávány metodou GET protokolu HTTP. Prvním povinným parametrem je znění původního dotazu na Elasticsearch, díky tomuto parametru výsledky funkce korespondují s kontextem aktuálního vyhledávání v portálu. Dalším parametrem je pak klíčové slovo nebo jeho část, které se hledá v dostupných hodnotách fasety.

Po úspěšném dotazu funkce vrací strukturu JSON ve stejném tvaru jako například v **checkedOptions** zmiňovaném v kapitole Vyhledávání

3.2.8 Model fasety

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Kapitola 4

Experimenty a vyhodnocení

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.1 Testování na uživateli

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.2 Zátěžové testy

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.3 Rychlost obnovy dat

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.4 Možnost úprav

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla

tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Kapitola 5

Závěr

Cílem této práce bylo Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Řešení této práce mi bylo přínosné především rozšířením obzorů v oblasti práce s neregulární databází a také využitím moderní technologie Vue.js.

5.1 Možnosti pokračování

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Literatura

- [1] Code Conquest: *Introduction to Web Development*. ©2019, [online; navštíveno 27.04.2019].
URL <https://www.codeconquest.com/what-is-coding/web-programming/>
- [2] European Union: *About CORDIS*. ©2019, [online; navštíveno 27.04.2019].
URL <https://cordis.europa.eu/about/>
- [3] Liferay Inc.: *What is a Web Portal?* ©2019, [online; navštíveno 20.04.2019].
URL <https://www.liferay.com/resources/l/web-portal>
- [4] Roy Williams: *Python 3 at Lyft*. Youtube, 2018-01-10, [online; navštíveno 23.04.2019].
URL <https://www.youtube.com/watch?v=aetoXWRt24k&t=164>
- [5] STANĚK, P.: *Automaticky aktualizovaný webový portál*. Bakalářská práce, Vysoké učení technické v Brně, Fakulta informačních technologií, Brno, 2016, vedoucí práce Smrž Pavel.
URL <https://www.fit.vutbr.cz/study/DP/BP.php?id=18568&y=2015>
- [6] Steve Cohen: *What challenges has Pinterest encountered with Flask?* Quora - A place to share knowledge and better understand the world, 2015-01-08, [příspěvek v diskuzním fóru; navštíveno 23.04.2019].
URL <https://www.quora.com/What-challenges-has-Pinterest-encountered-with-Flask/answer/Steve-Cohen?share=1&srid=hXZd>
- [7] UITS Knowledge Management team: *What is a Web Portal?* Indiana University Knowledge Base, 2018-01-08, [online; navštíveno 20.04.2019].
URL <https://kb.iu.edu/d/ajbd>