



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

David Nápravník

Softwarové řešení digitálních archivů

Katedra teoretické informatiky a matematické logiky

Vedoucí bakalářské práce: Mgr. Macková Kateřina

Studijní program: Informatika (B1801)

Studijní obor: IPSS (1801R048)

Praha 2021

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

TODO Poděkování:
Petra Hoffmannová
Kateřina Macková
Anička Yaghobová
Monika Bošániová

Název práce: Softwarové řešení digitálních archivů

Autor: David Nápravník

Katedra: Katedra teoretické informatiky a matematické logiky

Vedoucí bakalářské práce: Mgr. Macková Kateřina, Ústav formální a aplikované lingvistiky

Abstrakt: Za použití moderních webových a databázových technologií zmodernizujeme zastaralé knihovní systémy. Snížíme nároky na výkon serveru a síťového rozhraní. Uživateli těchto systémů poskytneme intuitivní a rychlou webovou aplikaci. Návštěvníkovi pak poskytneme nejen onen knihovní systém, ale celý balíček různých objektů, od novinek až po 3D modely historických objektu. Vývojáři a automatizované systémy budou mít přístup k API, jež bude poskytovat uchovávaná data. Zároveň bude sloužit i jako prostředek při komunikaci modulů systému s databází, např. pro umožnění vyhledávání pomocí neuronové sítě.

Klíčová slova: digitální archiv, web, databáze

Title: Software solution for digital archives

Author: David Nápravník

Department: Department of Theoretical Computer Science and Mathematical Logic

Supervisor: Mgr. Macková Kateřina, Institute of Formal and Applied Linguistics

Abstract: TODO Abstrakt en

Keywords: digital archive web database

Obsah

1. Úvod

Cílem této práce je vytvoření webové aplikace pro uchování a následné poskytování informací o historických zdrojích, především z oblasti Krkonoš.

1.1 Motivace

Cílem tohoto projektu je vytvoření webového rozhraní pro ukládání a zobrazování historických i současných záznamů z oblasti Krkonoš.

2. Analýza a požadavky

Vybíráme jen z open source produktů, abychom mohli nahlédnout do jejich kódu a lépe porozumět implementaci jejich komponent. Lépe se pro takový systém vyvíjejí nezávislé moduly a obvykle mají o dost větší komunitu vývojářů a přispěvatelů.

2.1 Požadavky na systém

Systém bude přístupný jako webová aplikace skrze moderní webové prohlížeče. Zadavatel bude moci přidávat, prohlížet, měnit a mazat metadata a k nim příslušné indexy. Redaktor bude moci přidávat a měnit články a novinky na webu. Uživatel bude moci vyhledat záznam podle několika různých kritérií a poté jej zobrazit, též bude moci zobrazit příspěvky na webu, včetně novinek. Externí programátoři budou moci k systému přistupovat přes API a získávat nebo měnit data.

2.2 Pro koho je systém určen

Systém bude sloužit pro historiky jako úložiště dat a pro veřejnost jako jejich zdroj.

2.3 Existující produkty

2.3.1 KOHA



url: <http://www.koha.cz/>

Koha je nejrozšířenější open source systém s širokou komunitou. Byla vyvinuta na Novém Zélandu roku 2000. Ale je stále udržovaná a stále rozšiřovaná (nejnovější update je z konce roku 2020)

Používá SQL databázi. Je psaná v Perlu, na frontendu využívající javascript (ale není jej tolik).

2.3.2 Evergreen



url: <https://eg-wiki.osvobozena-knihovna.cz>

Další z řady knihovních systému. Používán např. Pedagogickou a Teologickou fakultou v Praze.

Používá sql databázi, vykreslování probíhá na serveru a nemá uživatelsky přívětivé prostředí.

2.3.3 SLIMS



url: <https://slims.web.id/>

Systém se základní funkcionalitou a přívětivým vzhledem. Není v češtině. Není primárně určen jako knihovní systém, spíš je to univerzální systém na cokoliv, proto není až tak efektivní a jeho nastavování by zabralo mnoho času.

2.3.4 Souhrn

Evergreen a SLIMS jsou systémy, které potřebují silně vyškolenou osobu, aby se o systém starala, narozdíl od KOHY, která je intuitivnější a pro nové uživatele přívětivější. Při zachování stejné, možná i lepší funkcionality.

3. Návrh projektu

3.1 Výběr technologií

3.1.1 Frontend

Trendy v oblasti vývoje webových aplikací se mění rychle avšak velkou změnou, která ovlivnila celý způsob principu jak na stránku nahlížet přinesla technologie **Single page application**, jež je implementovaná např. v knihovně **React** zaštiťovanou společností Facebook.

Single page application

Single page application je technologie umožňující vykreslení jiné stránky, bez nutnosti posílání requestu na server. Uživatel si při prvním spuštění webu stáhne celý balíček webu a při opětovném načtení většinou sahá jen do své lokální cache. Javascriptová knihovna (v tomto případě React) poté stránku překresluje při uživatelské interakci. V případě nutnosti stažení / posílání dat mezi serverem a uživatelem (např. editace záznamu, nebo načtení existujícího záznamu) se volá pouze request k API webové služby a tělo requestu obsahuje pouze užitečné (ne-redundantní) informace.

React

Knihovna React poskytuje single page application technologii. Jedná se o dobře udržovanou knihovnu, jež byla vyvinuta Facebookem, jakožto náhrada zastaralého konceptu renderování stránky na serveru. Díky tomu servery nemusejí ztrácet výkon s každou změnou na stránce a výkon k renderování se bere z PC uživatele. Jádro této knihovny je velmi dobře optimalizované a poskytuje i řadu debutování nástrojů, což je pro větší projekty nepostradatelná výhoda.

Dalsi mozne technologie

Velmi často vykreslování stránek probíhá na serveru, se systémy jako jsou WordPress, psaný v PHP. Takovýto systém je velmi dobře uživatelsky přívětivý, ale z pohledu výkonu má velmi obrovský overhead. V případě implementace knihovničního systému by to znamenalo vykreslovat celou stránku (hlavičku, tělo i zápatí) na serveru, na druhé straně single page application nic nerenderuje, pouze pošle požadované informace.

3.1.2 Backend

Mít single page aplikaci na frontendu znamená, že na backendu musí existovat API, od kterého bude frontend čerpat data. Navíc zde potřebujeme i systém pro statické odesílání balíku celé webové stránky. V rámci udržitelnosti jsem se rozhodl využít jazyk Javascript stejný jako pro frontend. Express.js je knihovna která umožňuje komplexní správu requestu a stala se tudíž jasnou volbou.

Express.js

Express.js poskytuje odesílání statických stránek (Reactího balíku v našem případě), custom requesty pro rozmanité API a také odesílání a lokální ukládání statických souborů, jako obrázků, wordovských i pdf dokumentů atd.

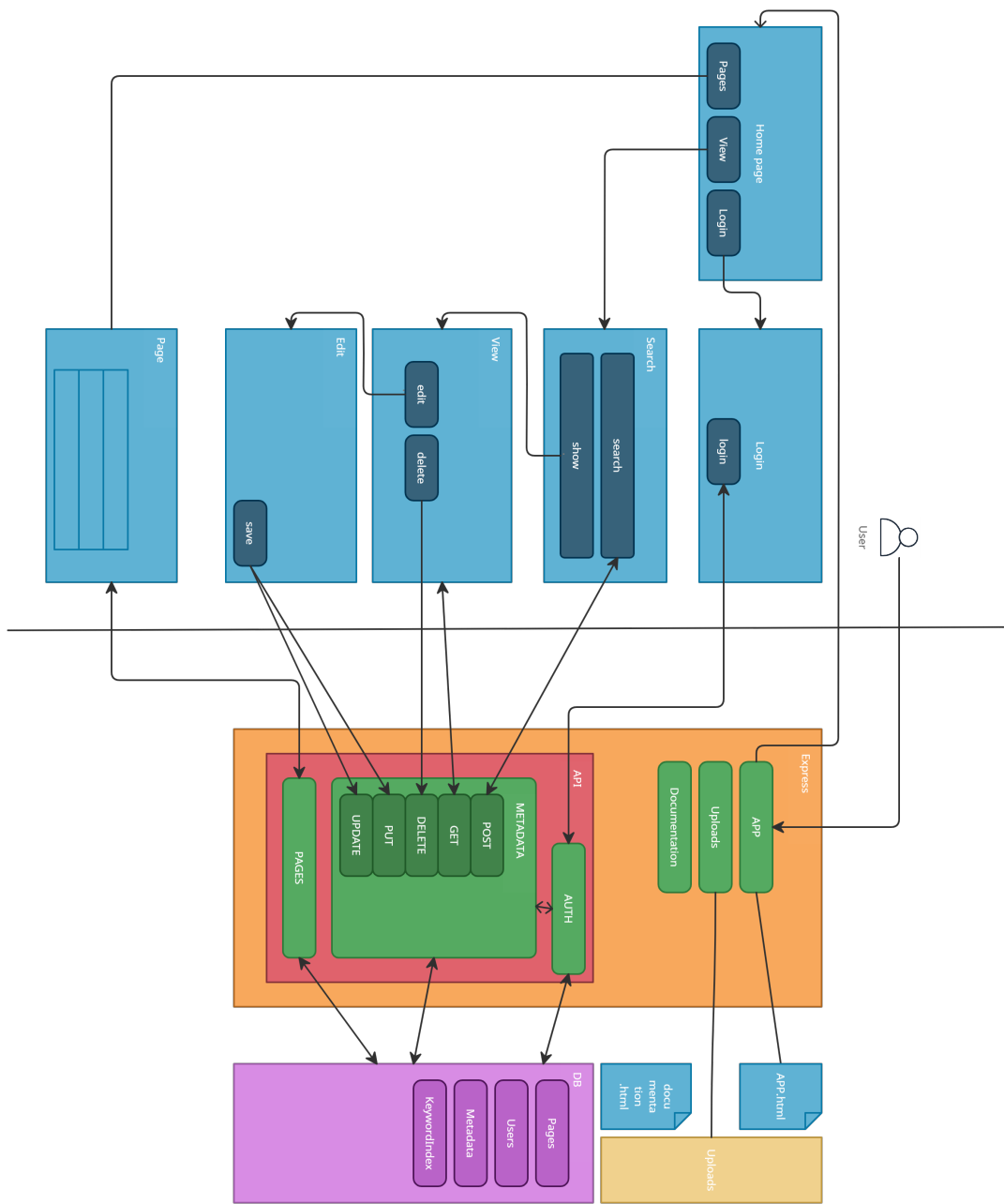
MongoDB

MongoDB je databázový systém typu non-sql. Což primárně znamená, že data neuchovává v tabulkách, ale v tkz. schématech. Což má mnoho výhod, největší je, že nekompletní záznamy nezabírají svými nevyplněnými daty místo v DB a ukládá se opravdu jen to co je potřeba. Další výhodou, je styl ukládání dat a komunikace s DB. Databáze si data uchovává ve formátu BSON (binární JSON rozšířený o datové typy). O data si aplikace žádá pomocí query, která je zcela odlišná od těch u sql-like databází, primárně se zde neposílá query ve formátu string ale jako JSON objekt, díky čemuž např. nenastane známa SQL injection. Znovu ve formátu JSON poté data vrátí aplikaci.

Další možné technologie

Díky oddělení frontendu a backendu (narozdíl od např. WordPressu) je možné na backend nasadit téměř cokoli co umí posílat requesty. Příkladem tomu můžou být skripty v jazycích PHP, C#, Python, nebo Perl. Ale vzhledem k tomu, že jedním z modulů bude neuronová síť na pokročilé vyhledávání, vybíral jsem mezi Pythonem a JavaScriptem, jakožto 2mi jazyky, které mají velmi dobré knihovny pro práci s neuronovými sítěmi.

3.2 Diagram systému



4. Implementace backendu

Backend je část softwaru, která je umístěna na serveru a uživatel z ní vidí jen výstup, jež dostane. Slouží pro přijímání požadavku od klienta a odesílání případné odpovědi nebo provedení nějaké činnosti bez výstupu.

Součástí backendu by mělo být API, které slouží jako toto spojení mezi serverem a klientem. Dále často obsahuje CRON tabulku, která periodicky vykonává nějaký script.

4.1 Server

Stroj na kterém běží aktuální verze systému běží pod systémem Linux (přesněji Ubuntu). Na virtual machine poskytnuté matematicko-fyzikální fakultou UK. Databáze je pak umístěna na serveru u Googlu, kvůli snazší konfiguraci. Není však problém kdykoliv tuto DB přesunout na stejný stroj, na kterém běží zbytek backendu a snížit tím prodlevu způsobenou prací s databází.

4.2 Knihovny

Pro rychlý začátek vývoje byla použita knihovna Express.js, která umožňuje práci s http requestsy potřebnými pro fungování API a to bez většího množství konfigurace ze začátku.

Pro práci s databází byla použita knihovna mongoose, která po rychlé konfiguraci umožňuje práci s databází typu MongoDB.

Dalšími menšími pomocnými knihovnami, jsou **cors** (Cross-Origin Resource Sharing) napomáhající s nastavením hlavičky u API requestu, **md5** pro šifrování hesel a hašování dat a nakonec **cookie-parser** zjednodušující práci s cookies.

4.2.1 Knihovna express.js

Jedná se o minimalistickou a zároveň velmi silnou knihovnu poskytující všestranné prostředky pro web. Obsahuje funkce pro jednoduchou správu HTTP metod a díky tomu je vytváření i větších API jednoduché a přehledné. Má velmi dobrý výkon, který se sice nedá srovnávat s knihovnami psanými v jazyce C, ale z JS knihoven je jeden z nejrychlejších.

Nadstavba pro nahrávání souborů

Nadstavba **express-fileupload** umožňuje v těle requestu rozeznat a zpracovat soubor, který se následně pomocí knihovny **fs** ukládá na server, specificky do složky uploads.

4.2.2 Knihovna mongoose

Ideální knihovna pro práci s databází typu MongoDB. Poskytuje funkce pro snadné připojení k databázi a komunikaci s ní. Hlavní výhodou této knihovny jsou modely, validace a vestavěné přetypování (javascript je dynamicky typovaný

jazyk). S modely se pracuje pomocí tkz. **promise**, která umožňuje řadit akce za sebe, ve stylu:

```
Model.find(body)
  .limit(_limit || 5)
  .exec()
  .then(result => { res.status(200).json(result) })
  .catch(err => { res.status(500).json("something went wrong") })
```

4.3 Dokumentace API

Na adrese <http://quest.ms.mff.cuni.cz/prak/api/documentation> se nachází statický soubor s dokumentací. Celá stránka je zapouzdřena do jediného souboru, který se po načtení vykreslí na straně uživatele, tudíž nezatěžuje server, ale především se dá stáhnout a prohlížet offline.

4.3.1 Vykreslovací engine pro dokumentaci

Aby bylo možné vykreslit stránku až na straně uživatele, je nutné přenášet i script, který to obstará. Tento script se skládá ze dvou částí. Engine na vykreslení a data samotná, která jsou uložena ve formátu JSON.

Script projde veškerá data a podle typu a kontextu postupně vytváří html elementy podle vestavěných šablon a přidává jim příslušné styly a ovládací prvky.

4.4 Routes

Pro rozpoznání, která akce se má vykonat při různých dotazech, se porovnává jak adresa dotazu tak i metoda. Nejdříve se vezme v potaz cesta dotazu za statickou předponou „<http://quest.ms.mff.cuni.cz/prak/api/...>“. Jakmile máme vubranou cestu zjistí se, co vlastně dotaz potřebuje udělat a to jednou z těchto metod:

- POST - většinou obecný dotaz s query v těle
- GET - dotaz požadující 1 záznam, většinou podle ID
- PUT - vytvoření nového záznamu
- PATCH - změna v existujícím záznamu
- DELETE - smazání záznamu

Pokud uživatel má požadované oprávnění, akce se provede. V každém případě uživateli přijde zpětná vazba o úspěchu resp. neúspěchu dotazu. Tělo této zprávy může obsahovat požadovaná data, potvrzení o úspěchu, nebo i chybová zpráva a její příčina.

4.4.1 Uživatel

Slouží pro správu uživatelských účtů. Zakládání nového účtu může zavolat kdokoliv, avšak na zbylé akce, jako mazání, nebo změnu hesla má právo pouze majitel a administrátor. Stejně tak heslo a sessionID nejsou přístupné zvenčí.

4.4.2 Autorizace

Pro ověření, zda přihlášený uživatel má příslušná práva (a to na straně serveru, jelikož lokálně si je může libovonně upravit) se používá **sessionID**, které se poté porovnává v databázi s uživatelem a jeho skutečnými právy.

SessionID získáme po odeslání korektních přihlašovacích údajů. Případně jej ztratíme při odhlášení, nebo expiraci (která je aktuálně nastavena na 1 rok).

4.4.3 Záznam

Pro prohlížení záznamu, resp. jejich vyhledávání není třeba žádné oprávnění. Avšak pro jejich editaci resp. mazání je potřeba mít přiřazena práva pro zápis.

4.4.4 Stránka

Zobrazení stránky též nevyžaduje žádná práva, ale k jejich vytváření je třeba mít roli tkz. „CMS editor“, která jak název napovídá definuje uživatele jakožto editora článků a příspěvků.

4.4.5 Nahrávání souborů

Pro nahrávání souborů je třeba práv pro zápis, stejně jako pro mazání. Zobrazení pak žádná práva nevyžaduje, ani nevyžaduje přístup ze stejné domény.

4.5 Modely

Model nebo též schéma je popis záznamu v databázi, obsahuje datový typ, formát a může obsahovat i referenční cestu. Vpodstatě se jedná o převodní tabulku, aby se Javascript a MongoDB schodli na datovém typu a struktuře (především pro případ reference, nebo pole dat). Část která je pro uživatele nejvíce viditelná je požadavek na unikátní hodnotu pole, nebo požadavek aby hodnota byla nenulová.

4.5.1 Záznam

Každý záznam má přesný popis v online dokumentaci API v pravém sloupci, kde je pravidelně aktualizován.

4.5.2 Stránka

Záznam stránky obsahuje název, jazyk, krátký popis, kategorii a obsah samotný. Pro možnost sledování změn je zde i automaticky generovaný seznam editorů a časů jejich editace.

4.5.3 Uživatel

Každý uživatel se přihlašuje emailem a heslem, s tím že heslo není uloženov tkz. raw formátu, ale je výsledkem spojení hesla a soli (termín pro náhodný text pro zvýšení bezpečnosti proti bruteforce útoku) a výsledný hash (metodou md5) se uloží do databáze. Při přihlášení pak stačí porovna hash soli a hesla s údaji v databázi. Pro ověřování práv ale neslouží heslo, ale sessionID, které je unikátní pro každého uživatele a časem nebo manuálně může expirovat a je nutné se přihlásit znovu. V záznamu uživatele je uloženo i jméno a příjmení pro případ, že by bylo potřeba zjistit kdo např. psal jaký článek a zároveň nebyl prozrazen email. Každý uživatel má určitá práva a to v libovolné kombinaci z:

- read - právo pro čtení, nahlížení do záznamů a jejich vyhledávání (toto právo má i nepřihlášený uživatel).
- write - právo pro zápis v rejstříkách, umožňuje editovat, vytvářet a mazat záznamy. Dále má navíc právo ukládat na server obrázky a dokumenty.
- execute - administrátorské právo, umožňuje vytváření uživatelských účtů a změnu hesla uživatele, stejně jako jeho údaje.
- cms - práva pro editační změny co se týče obsahu stránek, jako jsou novinky a příspěvky.

5. Implementace frontendu

Frontend byl vyvíjen jako single-page application pomocí javascriptu a reactu. Navržen tak aby byl jednoduše ovladatelný přehledný a především velmi rychlý.

5.1 Server

Celá aplikace frontendu je uložena na serveru, včetně zdrojových kódů. Uživatel si ale stahuje pouze zkompilovanou aplikaci a případné externí zdroje.

5.1.1 Kompilace

Ačkoliv je javascript scriptovací jazyk a kompilaci provádí až za běhu, tak je možné jej zkompilovat předem. Při této kompilaci knihovny **webpack** a **babel** provedou několik zásadních kroků, z nichž nejdůležitějšími jsou projití kódu a jeho převedení na ECMAScript 5 (kvůli zpětné kompatibilitě), přetřídění a sloučení knihoven do jednoho zdrojového souboru (javascript se k uživateli dostane při prvním dotazu na server) a minifikace výsledného souboru, jež dokáže inteligentně projít kód a optimalizovat jeho textovou délku, čím výrazně sníží čas potřebný k načtení stránky.

5.1.2 NPM

Celá aplikace má strukturu balíčku NPM (node package manager), tudíž ji lze snadno spustit kdekoliv, kde je nainstalovaný nodejs a knihovna npm. Zároveň udržuje pomocí souboru **package.json** přehled o potřebných knihovnách (dependencies) a obsahuje též různé spouštěcí skripty, např. script pro build produkční verze aplikace.

5.2 Knihovny

Aplikace se skládá z více než tisíce knihoven a modulů, tudíž se podíváme pouze na ty největší a nejdůležitější z nich.

5.2.1 React

Tato JavaScriptová knihovna je vyvinuta společností Facebook a slouží pro tvorbu uživatelského rozhraní.

Její základem je vykreslování pomocí tzv. single-page a je vhodná zejména pro aplikace, kde se často mění data. Využívá pozměněnou JavaScriptovou syntaxi známou jako JSX (JavaScript XML), která umožňuje pracovat s HTML tagy uvnitř JavaScriptového kódu, bez nutnosti práce DOM objektem.

Výhodou single-page aplikace je pak i optimalizovaná práce s DOM objektem, který je bottleneck (úzkým hrdlem) při nesprávném použití (nebo lépe řečeno, při kterémkoliv použití, které něco s DOM objektem dělá).

5.2.2 Material-ui

Společnost Google (aktuálně Alphabet Inc.) se v roce 2014 rozhodla sjednotit svou grafickou podobu a vyvynula designový jazyk, jež se stal příručkou jak vytvářet uživatelsky přívětivé aplikace (nejen na webu, ale třeba i v aplikaci mobilním telefonu nebo tabletu). Základem tohoto stylu je realistická práce se světlem a uživatelskou interakcí. Vše odlehčené a intuitivní. Součástí tohoto grafického jazyka jsou i lehce čitelné fonty (Roboto), pochopitelné ikonky a systém barev a jejich kombinací.

5.2.3 i18n

Jelikož je projekt mířen i na nečeské uživatele, vyskytla se potřeba rozhraní překládat. Jednoduchým přístupem by bylo zkompileovat několik aplikací každou pro jiný jazyk, tudíž by se překlad odehrával na straně serveru. Bohužel to není neoptimálnější řešení a tudíž byla využita knihovna i18n, jež funguje na podobném principu jako single-page application a to sice, že dokáže měnit překlad stránky bez nutnosti stránku přenačíst a zároveň hlavní aplikace neobsahuje všechny jazyky při prvním načtení. Jazyky se postupně dostávají dle preference uživatele (vše samozřejmě na pozadí bez uživatelského zásahu). Velkou výhodou tohoto systému může být například situace, kdy je potřeba přeložit jeden nápis na stránce, případně sehnat anglický ekvivalent a nechceme přenačtením stránky přijít o již vyplněná data.

5.2.4 Babel

Jak už bylo zmíněno výše, aplikace se kompiluje a za tuto část je zodpovědná právě knihovna Babel.

Hlavní výhodou je kompilace kódu v ES6+ (EcmaScript v6, neboli Javascript v6) do ES5 (EcmaScript v5, neboli Javascript v5) Tímto převodem získáme zpětnou kompatibilitu pro starší javascriptové enginy.

5.2.5 Webpack

Aby se z tak obrovského množství knihoven a souborů stal jediný spustitelný soubor pomáhá knihovna Webpack, která má za úkol kód zkomprimovat a sjednotit. Touto optimalizací si ušetříme množství dotazů, jež bude muset náš server přijmout (a ekvivalentně s tím uživatel odeslat).

Zároveň nám umožňuje používat formát **SCSS**, což je nadstavba nad klasickým CSS, jež rozpoznávají prohlížeče a dovoluje nám takto vytvářet **stylovací moduly**.

5.3 Rozhraní

Jelikož se jedná o složitější systém a všechno nemůže být naházené v jediném souboru, tak je použita hierarchie která vypadá následovně:

- Hlavní soubor celého webu - provádí routing a obaluje celou aplikaci pomocnými wrapery (např. i18n překladač, Mui pro jednotný styl, Snackbar-Provider pro vyskakovací toasty a CookiesProvider pro jednotný přístup ke cookies)
- Scény - Stránky diametrálních vlastností a funkcionalit
- Komponenty - malé na sobě nezávislé „černé krabičky“ poskytující určitou funkcionalitu

5.3.1 Scény

Admin

Administrační rozhraní je přístupné pouze s oprávněním „execute“. Zde může administrátor nebo správce uživatelů vytvářet nové účty nebo konfigurovat stávající.

Jelikož se hesla v databázi ukládají zahešovaná, tak je nelze zobrazit ani adminovi. Dala by se přidat funkcionalita pro jejich crackování, aby se našli uživatelé se slabými hesly, ale to by bylo zbytečné plýtvání serverovými prostředky.

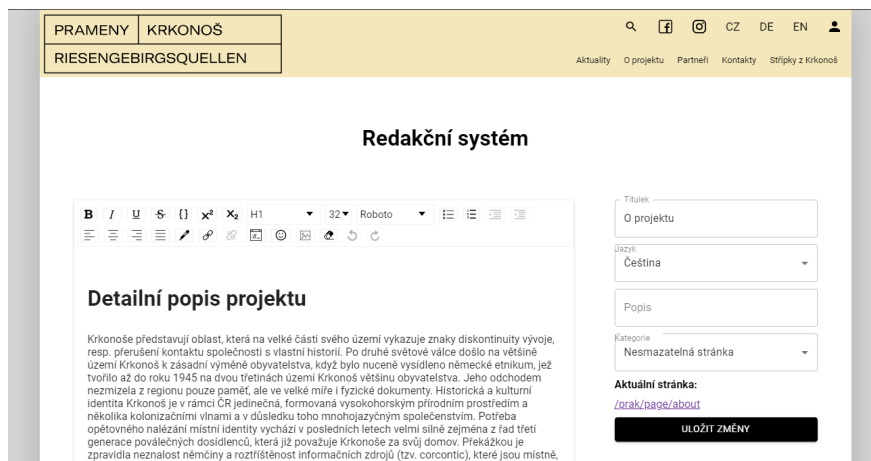
Jméno	Příjmení	Email	Heslo	Čist	Vkládat	Spravovat (Admin)	CMS (Redakční systém)	ULOŽIT ZMĚNY	SMAZAT UŽIVATELE
admin	Příjmení	Cenzurováno	*****	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	ULOŽIT ZMĚNY	SMAZAT UŽIVATELE
Kateřina	Brožová	Cenzurováno	*****	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ULOŽIT ZMĚNY	SMAZAT UŽIVATELE
David	Babka	Cenzurováno	*****	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	ULOŽIT ZMĚNY	SMAZAT UŽIVATELE
Olga	Hájková	Cenzurováno	*****	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	ULOŽIT ZMĚNY	SMAZAT UŽIVATELE
Václava	Horáková	Cenzurováno	*****	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ULOŽIT ZMĚNY	SMAZAT UŽIVATELE

CMS

Vytváření a editace obsahu stránek se provádí v redakčním systému (Content Management System, zkráceně CMS). Do něj se uživatel dostane pouze po přihlášení a s platnými právy pro redakční systém.

Úpravy redaktor provádí v prostředí WYSIWYG (What you see is what you get, neboli co vidíš to dostaneš). V pravém panelu pak zadá název, případně krátký popis, který se hodí např. pro aktuality a kategorii.

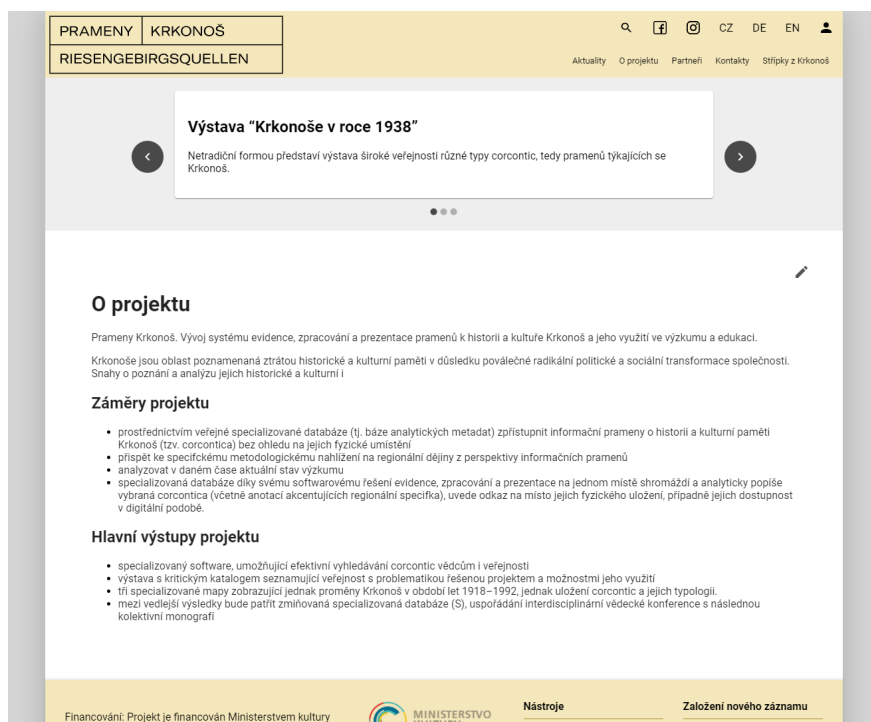
WYSIWYG editro podporuje velkou škálu stylování a formátování. Velkou výhodou je možnost nahrávání obrázků, kdy se po přetažení automaticky uloží na server do složky uploads a jejich odkaz se vloží do stránky.



Domovská stránka

Výchozí stránka na kterou se uživatel dostane, pokud nezadá přesnější cestu v url, je domovská stránka.

Na této stránce najde dynamické zobrazování novinek a krátkého popisu v horním panelu a stránku s názvem „homepage“, což je obyčejná stránka editovatelná v redakčním systému a dovoluje tedy redaktorům měnit i domovskou stránku.



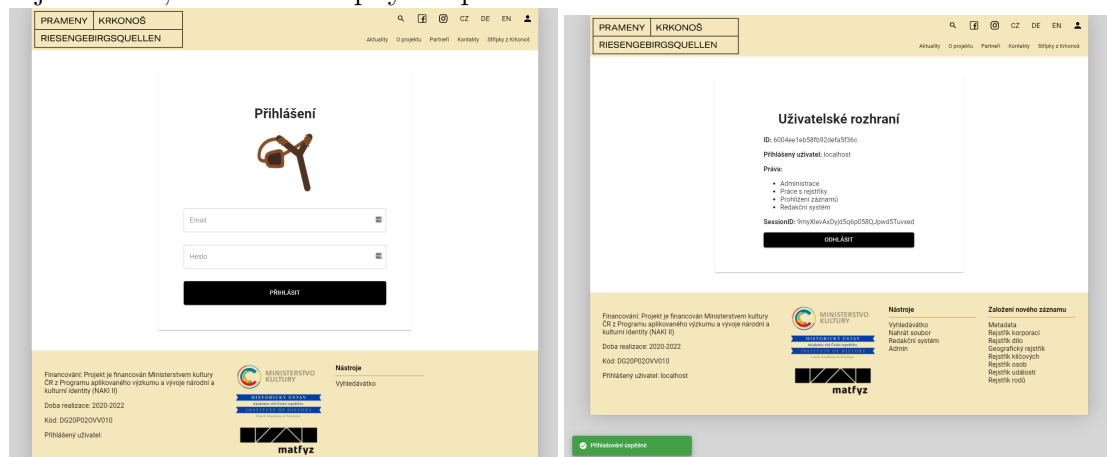
Zobrazení stránky

Každá stránka vytvořená v redakčním systému je přístupná na adrese „.../prak/page/#jmeno“. Zde si ji můžou uživatelé zobrazit. Pokud má uživatel i právo pro redakční systém, tak se mu v pravém horním rohu zobrazí ikonka pro editaci, díky které se snadno dostane do redakčního systému pro danou stránku.



Přihlašování

Pokud uživatel klikne na ikonku postavičky vpravo v hlavičce, nebo se pokusí dostat na stránku, ke které nemá oprávnění, je přesměrován do přihlašovacího prostředí. Pokud se poté přihlásí, nebo již přihlášen byl zobrazí se mu podrobnosti o jeho účtu, včetně dostupných oprávnění.



Kontaktní formulář

Pro zpětnou vazbu, či otázky k projektu, můžou návštěvníci využít kontaktní formulář.

PRAMENY

KRKONOŠ

RIESENGBIRGSQUELLEN

f

@

CZ DE EN

Aktuality

O projektu

Partneři

Kontakty

Stípký z Krkonoš

Kontakty

Máte-li otázku, podnět či připomínku, kontaktujte nás prosím prostřednictvím kontaktního formuláře.

Jméno *

Email *

Zpráva *


ODESLAT

Kontaktovat můžete také koordinátora projektu kohoutek@hiu.cas.cz či jednotlivé řešitele.

Financování: Projekt je financován Ministerstvem kultury ČR z Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI II)

Doba realizace: 2020-2022

Kód: DG20P020VV010



MINISTERSTVO KULTURY

HISTORICKÝ ÚSTAV

Ústav pro českou minulost

ÚSTAV PRO ČESKOU MINULOST

Nástroje

Vyhledávátko

Vyhledávátko

Pomocí ikonky lupy v hlavičce se dostaneme do vyhledávacího prostředí. Po výběru kde chceme hledat se nám načtou políčka k vyplnění. Uživatel do nich může zadat libovolný výraz a to s podporou regexp syntaxe. Ihned při vyplňování se načítá až 5 prozatím nejvhodnějších výskytů hledání. Při kliknutí na tlačítko „vyhledat“ se provede vyhledání všech odpovídajících záznamů. Nalezené záznamy se objevují v tabulce pod vyhledávacím polem. Tabulka dokáže záznamy třídit po kliknutí na hlavičku příslušného sloupce. Po kliknutí na záznam se otevře editační prostředí pro vybraný záznam.

PRAMENY

KRKONOŠ

RIESENGBIRGSQUELLEN

CZ

DE

EN

Aktuality

O projektu

Partnery

Kontakty

Střipky z Krkonos

Vyhledavátko

Metadata

VYHLEDAT

Hlavní autor (Rejstřík osob)

Hlavní název

...

Datum vydání nebo vzniku

Nápověda

Vyhledání všech záznamů: nechte všechna pole prázdná a klikněte na Vyhledat

Interaktivní režim: už během zadávání se ukáže 5 nejlepších shod

Zobrazení záznamu: klikněte kamkoliv na příslušný záznam v tabulce níže

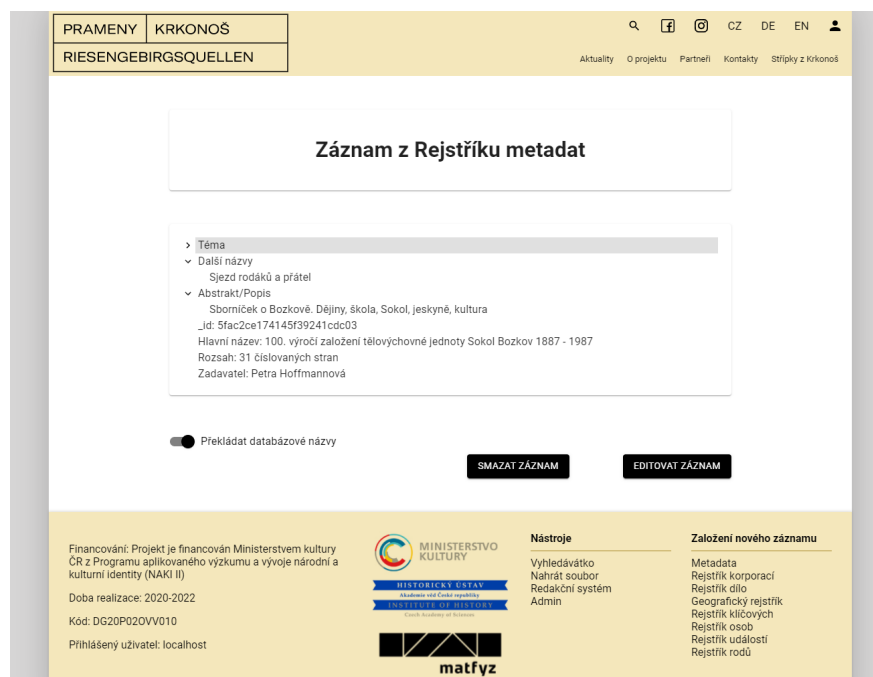
Hlavní autor	Hlavní název	Datum vydání nebo ...
	Paměti a folklorní sběry Jaroslava Večerníka z...	1950-1999
	Skica pro obec Horní Albeřice	
	"Tondova soukromá sbírka"	
	100. výročí založení tělovýchovné jednoty Sok...	
	90 let Církve československé husitské	

1-5 of 5

Zobrazovávátko

Záznam uložený v databázi si je možné zobrazit přes toto rozhraní poté co se k němu dostanete přes vyhledávací rozhraní, nebo pomocí unikátního url linku, který se po vytvoření záznamu již nemění.

Po načtení se uživateli zobrazí záznam, tak jak je uložen v databázi, případně s přeloženými názvy, pokud to má pomocí spodního přepínačka povolené. Dole též najde dvě tlačítka pro smazání záznamu a editaci, jež uživatele přesune do editačního rozhraní s aktuálním záznamem (samozřejmě jen pokud má dostatečná oprávnění).



Upravovátko

V tomto prostředí je možné záznam upravovat a jedná se o nejkomplikovanější scénu ze všech.

Každý typ záznamu a každý podtyp metadatové struktury má vlastní příslušná datová pole. Ty které umožňují mít více hodnot mají pravo u sebe tlačítka plus a mínus, jehož velikost určuje velikost bloku, jež je jedním celkem. Plusové tlačítko přidává další datové pole, zatímco mínus datové pole vyprázdní a poté odebere z náhledu.

Datová pole jsou roztržena do menších podbloků, které se nechají minimalizovat (bez ztráty dat), aby uživateli usnadnily orientaci. Po minimalizaci resp. maximalizaci bloku může dojít k převážení sloupců a některé bloky tak mohou střídát levý a pravý sloupec, tak aby celková stránka byla co nejkratší a prázdného místa byla co nejméně.

Některé záznamy mají možnost nahrát přílohu, ta se aktivuje po kliknutí na tlačítko uploadu, s ikonkou šipky směrem nahoru. Nahraný soubor se ihned přenáší na server a do příslušného pole automaticky zaznamená aktuální url nahrávaného souboru.

Po kliknutí nahrát se buď objeví červené vyskakovací okno s chybou, proč záznam nelze uložit (většinou zapříčiněný zadáním unikátního názvu, jež již je v databázi uložen). Nebo se stránka znovu přepne do zobrazovátka a vyskočí zelené potvrzovací okénko.

5.3.2 Komponenty

Komponenty jsou vlastně zapouzdřené třídy, které se dají použít vícekrát. Každá komponenta je chodná pro React systém a většinou je psaná bez tkz. hooků.

ComboBox

ComboBox je textové pole, které při vyplňování realtime vyhledává v databázi vyhovující shody, které zobrazuje ve formě nabídky pod tímto textovým vstupem. Po kliknutí na nabízenou položku se do pole vyplní hodnota a na pozadí se do pole uloží ID záznamu (to se poté odesílá na server). Dokud uživatel některý ze záznamů nevybere, pole není v konzistentním stavu a nemá možnost se odeslat na server, což značí ikonka ✖, v případě že pole má přiřazené id záznamu a zobrazuje název, rozsvítí se ikonka ✓, která značí že data z tohoto pole se při uložení správně odešlou na server. Popisky často bývají dlouhé a proto je zde i miniaturní posuvník, se kterým se dá pohodlně pracovat pomocí najetí myši na políčko, přidržení klávesy *Shift* a točením kolečka na myši.

Zápatí

Jako každá správná stránka i v tomto projektu najdeme zápatí, jež obsahuje povinné údaje o projektu a rychlé odkazy. Jelikož se jedná o projekt Evropské unie, nalezneme zde povinné údaje, kód

projektu a loga sponzora, zadavatelů a vývojářů. Níže je ukazatel aktuálně přihlášeného uživatele. V pravé části jsou pak rychlé odkazy, které se zobrazují v závislosti na právech přihlášeného uživatele. Nepřihlášený uživatel tudíž v této části uvidí pouze jedno menu „Nástroje“ s jedinou položkou „Vyhledávátko“ a po přihlášení se viz obázek menu rozroste o další rychlé odkazy.



Navigační menu

V navigačním menu nalezneme rychlé odkazy na vyhledávátko (tlačítko s ikonkou lupy), facebook a instagram projektu Prameny Krkonoš a přihlašovací stránku. Tlačítka pro přepínání mezi třemi jazyky. Níže pak hlavní stránky a seznamové stránky (např. stránka s aktualitami).



Textové pole s validací

Jako kontrola, že uživatel zadává data ve srozumitelném formátu (tak aby je dokázal pochopit další uživatel nebo aby je databáze správně interpretovala) jsou data kontrolována na straně klienta. V případě že data požadovaný formát nemají, pod políčkem se zobrazí varovný nápis a záznam nepůjde uložit, dokud všechna pole nejsou ve správném formátu.

Souřadnice
N52.0 N53.0
Chybný formát souřadnic

Pole pro nahrávání souborů

Některé záznamy a stránky potřebují obsahovat přílohu ve formátu obrázku nebo dokumnetu.

Tyto soubory lze nahrát na server a do databáze uložit pouze odkaz na jejich umístění, protože databáze není stavěná na uchovávání obrázků a podobných relativně velkých souborů.

Příloha
/prak/uploads/tigg8.png

Indexy

Hlavní částí editačního rozhraní jsou indexové objekty, které určují, která pole se mají zobrazit při různých typech záznamů. Pro každý typ je zde JSON soubor, ve kterém je seznam polí a informace o nich. Mezi tyto informace patří

především popisem, struktura jak data odeslat do databáze, nápověda přístupná pomocí malého otazníčku a nepovinná pole jako požadavek na nenulovou hodnotu, nebo regexpový výraz pro kontrolu formátu dat.

5.3.3 Moduly

Funkcionalita, která není až tak často používána, se do hlavního souboru aplikace nepřidává a tudíž čistý frontend žádné moduly nepodporuje. Moduly jsou výhradně načítány ze serveru a neobsahují hlavní frontend.

5.4 Lokalizace

Překlad stránky do různých jazyků je řešen pomocí knihovny i18n, která je velmi efektivní a zároveň poskytuje snadný způsob jak texty překladu přidávat a editovat. Překlady navíc jsou rozloženy do více souborů a úrovní, takže se překlad velmi dobře škáluje i na větší projekty.

Výsledkem jsou lokalizační soubory ve formátu JSON, které frontend aplikuje na požadovaných místech. Pokud ovšem překlad chybí použije se nejbližší možná interpretace, podobný jazyk, nebo zdrojová adresa překladu a vývojáři ukáže zprávu o chybějícím překladu.

6. Provázání Backendu a Frontendu, API

6.1 Dokumentace online

Aktuální dokumentace, tak aby mohla být časem aktualizována a mohli do ní být připisovány další věci, se nachází na webu quest.ms.mff.cuni.cz/prak/api/documentation. Dokumentace je rozdělena na dvě části.

První část popisuje volání API. Každá metoda (GET, POST atd.) má svůj vlastní účel jež je popsán uvnitř url a další možné parametry. Spolu s formátem requestu je zde i formát odpovědi. Podle kódu zjistíme, jestli byl náš požadavek úspěšný. Kódy jsou standardní podle "http status codes".

- **2xx** Všechno dopadlo dobře
- **4xx** Chyba je na straně klienta
- **5xx** Chyba je na straně serveru

V případě úspěchu (kód 200 - OK) se odešle odpověď na dotaz, nebo nic pokud request neměl za funkci něco vracet. V případě neúspěchu pak v odpovědi najdeme zprávu o chybě která nastala. Obvykle to u kódu 400 bývá odeslání duplicitního záznamu.

Druhá část popisuje strukturu schémat jednotlivých modelů. Jelikož databáze má datové formáty, zatímco JSON soubor ne (nebo alespoň ne tak rozsáhlé) musí i odesílaná data mít správný formát, nebo alespoň být validní po automatickém přetypování. Schéma tvoří JSON objekt jež toto schéma popisuje. Pokud je datový typ položky objekt, buď se opravdu jedná o objekt, nebo se jedná pouze o upřesnění datového typu. Klíčovými slovy jsou:

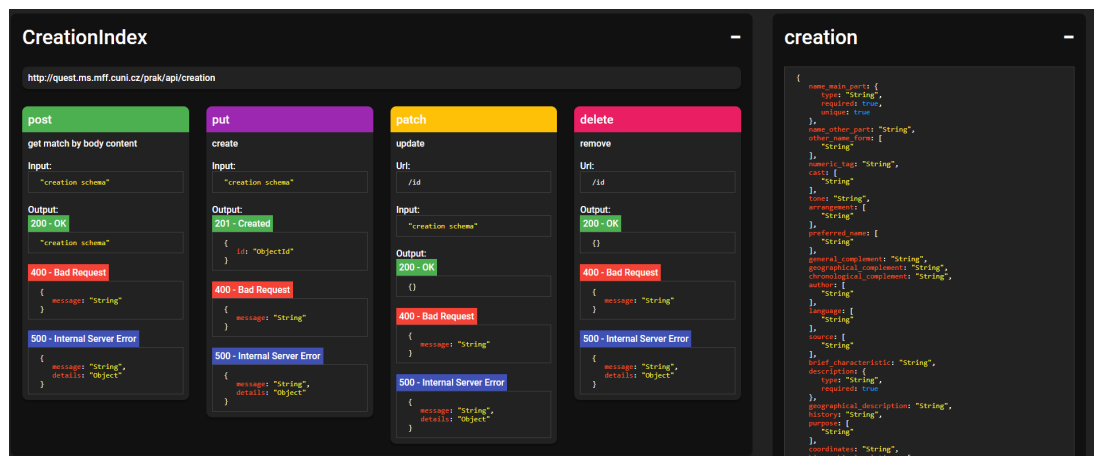
- **type**: datový typ
- **required**: true, pokud je tato položka povinná
- **unique**: true, pokud se zadaná hodnota nesmí schodovat s již existující položkou v DB
- **ref**: název schématu, na který se ID odkazuje
- **refPath**: speciální ref, umožňující uživateli zadat i název schématu, vůči kterému se odkazuje

Pokud je hodnota pouze string, pak je tato hodnota datovým typem.

6.1.1 Software pro online dokumentaci

Pro možnost stažení dokumentace a prohlížení offline, je vše zabaleno do jediného html souboru. Uvnitř je zdrojový kód programu, jež vykresluje stránku a zároveň data dokumentace.

Program vykresluje veškeré položky s daty dokumentace. Bloky zdrojových kódů jsou vysazeny fontem monospace a obarveny, aby uživateli poskytly rychlejší orientaci v kódu.



6.2 Backend

Na backendu je spuštěn express server (běžící pod nodejs), který zachytává requesty s url ...prak/api/... . podle cesty, jež je uvedena za /api se request předává příslušnému routeru. Výjimku tvoří adresa .../api/documentation, která rovnou přeposílá soubor s dokumentací a není tedy pro získání dokumentace třeba rozumět systému requestů hlouběji.

Při převzetí requestu jedním z mnoha routerů, se porovnává typ requestu (POST, PUT atd.) a případné detaily cesty v url. Při plném nalezení shody se provede ověření práv, pokud je to potřeba. Pokud request má požadovaná oprávnění je provedena příslušná funkce a uživateli je vrácena odpověď případně potvrzení o úspěchu. V případě že request nemá příslušná oprávnění, je vrácen kód 401. V případě že se na serveru něco pokazí vrátí se kód 400, nebo 500, podle typu problému.

6.3 API

API je přístupné na adrese quest.ms.mff.cuni.cz/prak/api/... .

6.3.1 Autentifikace

S každým requestem přichází v hlavičce i cookie, ten pro autentifikaci se jmenuje "sessionID". Podle něj se najde příslušný uživatel a porovnají se jeho práva a práva potřebná pro vykonání požadované funkce. Pokud jsou práva nedostatečná, vrátí se odpověď s kódem 401, v případě správného oprávnění, router vykoná

funkci jež danému requestu přísluší a pokud se nepokazí nic jiného, vrátí validní odpověď.

6.4 Frontend a volání API

Jelikož je celý projekt zamyšlen jako webová aplikace, nejstandardnější použití je volání API pomocí JS funkce *fetch()*, což je pouze technický alias pro XMLHttpRequest. Ale je možné jej volat jakkoliv jinak, dokud to bude validní http request.

6.4.1 Fetch

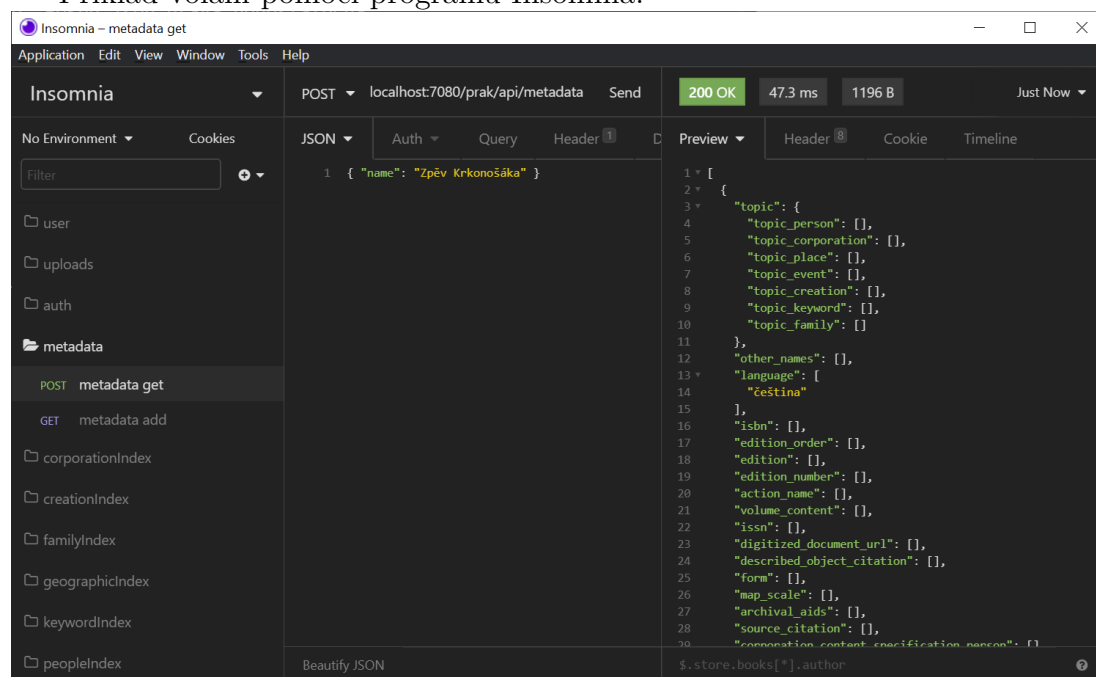
Dokumentace metody: mozilla

Příklad použití:

```
const url = "/prak/api/metadata"
fetch(url, {
  method: "POST",
  headers: { "Content-Type": "application/json" },
  body: JSON.stringify({ "name": "... " }),
})
.then(response => {
  if(!response.ok) throw response
  return response.json()
})
.then(response => {
  console.log(response)
})
.catch(error => {
  console.error(error)
})
```

6.4.2 Existující programy pro práci a testování API

Příklad volání pomoci programu Insomnia:



7. Moduly

7.1 Přidávání nových modulů

7.2 Aktuálně nasazené moduly

7.2.1 Modul hologram

Načítání modulu

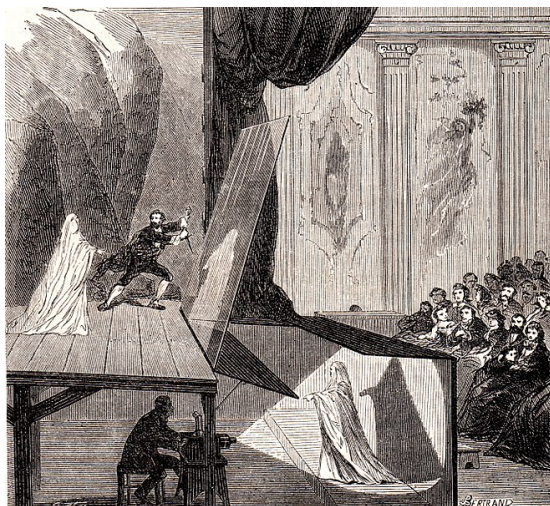
Tento modul obsahuje velkou knihovnu a tudíž není efektivní jej mít v primární aplikaci. Protože by se tyto knihovny načítaly, i když by uživatel s tímto modulem neměl v plánu pracovat. Což většinu času nebude chtít a tudíž by to pouze vedlo ke zpomalení systému. Systémem LAZY načítání je tedy celý modul odříznut a načítá se až explicitně při jeho použití.

Grafický engine

Pro vykreslování 3D modelu na webu slouží WebGL, grafická knihovna podobná staré verzi OpenGL. Reálně jsou dostupné dvě propracované knihovny pro práci ve WebGL. ThreeJS, které se zaměřuje na statické vykreslování a předvádění. BabylonJS, která naopak napodobuje známe programy jako unity a má podporu fyziky a dalších podknihoven užitečných pro tvorbu her. Pokud pomineme možnost naprogramovat vše od základu, nejvhodnější knihovnou je **ThreeJS**.

Z této knihovny využijeme moduly pro načítání modelu a textury, jejich spárování a efektu Peppers Ghost, který je implementován pomocí 4 nezávislých kamer.

Peppers Ghost Effect



source: Wikipedia

Jedná se divadelní trik, jež skládá pozorovateli dva obrazy přes sebe a vytváří iluzi hologramu nebo ducha. Původně byl vyvinut pro divadelní představení, v moderní době však dokáže simulovat i hologram, který známe ze sci-fi.

Modely a textury

Modely jsou z webu <https://3dwarehouse.sketchup.com> a textury z webu <https://www.textures.com/> a <https://texturehaven.com/>.

V programu blender pak byly tyto textury namapovány na objekt a bylo do nich zapéčeno ambient occlusion (stíny vytvořené stísněným prostorem).

8. instalace a spusteni

9. Reseni

9.1 výsledny web

9.2 uživatelska dokumentace

Závěr

Seznam obrázků

Seznam tabulek

Seznam použitých zkratek

A. Přílohy

A.1 První příloha