ZADÁNÍ BAKALÁŘSKÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Růžička** Jméno: **Tomáš** Osobní číslo: **484887**

Fakulta/ústav: **Fakulta informačních technologií**

Zadávající katedra/ústav: **Katedra softwarového inženýrství**

Studijní program: **Informatika**

Studijní obor: **Webové a softwarové inženýrství**

II. ÚDAJE K BAKALÁŘSKÉ PRÁCI

Název bakalářské práce:

Asterion - backend vizualizace časových os

Název bakalářské práce anglicky:

Asterion - timelines visualization backend

Pokyny pro vypracování:

Asterion je rozsáhlý svět s bohatou a komplikovanou historií. Cílem práce je vizualizovat historická data tohoto fiktivního světa pomocí vícera časových os vázaných na konkrétní téma, lokace, apod. Výsledný produkt bude běžet jako webová služba. Práce je součástí skupinového projektu Asterion.
1) Proveďte rešerši způsobů vizualizace historických dat.
2) Analyzujte a vhodně uspořádejte množinu časových dat světa Asterion poskytnutých zadavatelem.
3) Proveďte rešerši technologického řešení webové vizualizace.
3) Navrhněte prototyp - zaměřte se na backend a vytvoření vhodné databáze. Bezpečnost přenosu administrátorského přístupu. Možnosti privilegovaných uživatelů zadávat historická data.
4) Implementujte a nasaďte službu.
5) Důkladně otestujte bezpečnost a funkčnost služby.

Seznam doporučené literatury:

Jméno a pracoviště vedoucí(ho) bakalářské práce:

Ing. Radek Richtr, Ph.D., katedra softwarového inženýrství FIT

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) bakalářské práce:

Datum zadání bakalářské práce: **16.02.2021** Termín odevzdání bakalářské práce: _____

Platnost zadání bakalářské práce: _____

Ing. Radek Richtr, Ph.D. podpis vedoucí(ho) práce	podpis vedoucí(ho) ústavu/katedry	doc. RNDr. Ing. Marcel Jiřina, Ph.D. podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Student bere na vědomí, že je povinen vypracovat bakalářskou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v bakalářské práci.

. _____ _____

Datum převzetí zadání Podpis studenta

Bakalářská práce

ASTERION – BACKEND VIZUALIZACE ČASOVÝCH OS

Tomáš Růžička

Fakulta informačních technologií ČVUT v Praze
Katedra softwarového inženýrství
Vedoucí: Ing. Radek Richtr, Ph.D.
28. prosince 2021

Odkaz na tuto práci: Tomáš Růžička. *Asterion – backend vizualizace časových os.* Bakalářská práce. České vysoké učení technické v Praze, Fakulta informačních technologií, 2021.

Obsah

Seznam obrázků

Seznam tabulek

Seznam výpisů kódu

Chtěl bych poděkovat především Ing. Radku Richtrovi, Ph.D. za nehynoucí trpělivost s námi a za nadšení do projektu. Dále bych chtěl poděkovat Michaele Zimmermannové za spolupráci na projektu a všem, kteří si tuto práci budou číst, za jejich čas a pozornost.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principu při přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisu. V souladu s ust. § 2373 odst. 2 zákona č. 89/2012 Sb., občanský zákoník, ve znění pozdějších předpisu, tímto uděluji nevýhradní oprávnění (licenci) k užití této mojí práce, a to včetně všech počítačových programu, jež jsou její součástí či přílohou a veškeré jejich dokumentace (dále souhrnně jen „Dílo“), a to všem osobám, které si přejí Dílo užít. Tyto osoby jsou oprávněny Dílo užít jakýmkoli zpusobem, který nesnižuje hodnotu Díla a za jakýmkoli účelem (včetně užití k výdělečným účelum). Toto oprávnění je časově, teritoriálně i množstevně neomezené.

V Praze dne 28. prosince 2021 .....

Abstrakt

Asterion je fiktivní svět psaný hráči pen&paper her na hrdiny. Spisovatelé se musejí orientovat ve sledu událostí, pokud se chtějí vyhnout nesrovnalostem. Projekt spojený s touto prací se snaží vytvořit aplikaci pro tvorbu a vizualizaci událostí na časových osách. Cílem této práce je vytvořit API pro tuto aplikaci. API bylo implementováno pomocí Node.js, MySQL byla použita jako databáze a nginx byl použit jako webový server a pro šifrování pomocí ssl. Tento text obsahuje rešerši konkurenčních řešení, popis základních technologií a bezpečnostních principů a návrh aplikace. Aplikace byla otestována a shledána funkční a bezpečnou.

Klíčová slova asterion, časové osy, historické události, webová vizualizace, webová aplikace, backend, API, HTTP, cookies, autentizace, databáze

Abstract

Asterion is a fictional world written by players of a pen&paper role playing games. Writers have to navigate the web of events if they want to avoid inconsistencies. Project associated with this work is supposed to make an application for creation and visualization of these events and timelines. Goal of this work is to create a web API for such application. API has been implemented with Node.js, MySQL has been used as a database and nginx has been used as a web server taking care of directing and ssl certification. This document contains research of competing applications, description of basic technologies and security principles and a application design. Application has been tested and found to be functional and safe.

Keywords

asterion, timelines, historical events, web visualization, web application, backend, API, HTTP, cookies, authentication, database

Úvod

Co je Asterion

Tento projekt je určen pro hráče pen&paper hry na hrdiny zasazené ve světě Asterion a spisovatele, kteří tento svět pomáhají tvořit.

Hra je v podstatě diskuze hráčů o tom, co by se mělo ve světě odehrávat. Přičemž se dějové linky drží postav, které si vybrali, a nějakého základního příběhu, který jeden z hráčů nastínil.

Motivace

Jelikož Asterion je svět s velmi dlouhou historií a velkým množstvím hráčů, je složité vymýšlet příběh, ve kterém není rozpor s již zaznamenanými fakty a událostmi v tomto světě.

Hráči nemusí dokonce ani hrát právě v aktuální době. Mohou se rozhodnout, že budou hrát jako skupina pytláků někde uprostřed doposud zaznamenané historie světa. Poté je vhodné do příběhu zakomponovat již známé události (války, svatby, narození).

Tyto události jsou ale rozprostřeny v knihách, které spisovatelé napsali, a málo komu se chce číst tři tlusté knihy, aby zjistil, kdy se vlastně mlynář Pešek z Kominíkova stal mlynářem.

Cíl projektu a práce

Cílem projektu je vytvořit webovou aplikaci, která bude na časových osách jasně a přehledně vizualizovat události, které se odehráli v herním světě Asterion.

Tato aplikace by měla uživateli pomoci zmapovat sled událostí týkajících se nějakého tématu. Například pohyb několika postav najednou a zjistit tak, zda je možné, aby se náhodou potkaly v krčmě U Fredyho v pátek 13.

Cílem této práce je pak vytvořit technické zázemí pro tento projekt a databázi událostí a webové API, čili backend.

Postup

Nejprve byla provedena rešerše konkurentů a technologií. Následně proběhl návrh domény, databáze a základního API pro provoz vizualizace. To bylo poté také implementováno. Pak proběhla příprava testovacích nástrojů a testování. Dále byl proveden návrh a implementace části API odpovědná za autentizaci uživatelů a její důkladné otestování. Poté byla na řadě část API schopná přidávat, editovat a mazat data, tato část byla opět navržena, implementována a otestována na separátní instanci API a databáze. Nakonec proběhlo důkladné plošné otestování a finální nasazení.

Výsledky práce

Výsledkem práce je nasazená a spuštěná aplikace pro vyhledávání a editaci historických událostí ve formátu navrženém pro snadnou implementaci vizualizace dat na časových osách.

Seznam zkratek a pojmů

Zkratky

HTML	Hyper Text Markup Language
CSS	Cascading Style Sheets
JS	JavaScript
WASM	Web Assembly
HTTP	Hyper Text Transfer Protocol
HTTPS	HTTP Secure
Web API	Web Application Programming Interface
XSS	Cross-site scripting
CSRF	Cross-site request forgery
ACID	Atomicity, Consistency, Isolation, Durability
MVCC	Multiversion concurrency control

Pojmy

HTML	Jazyk určující strukturu webové stránky
CSS	Jazyk určující vzhled webové stránky
JS	Jazyk určující chování webové stránky
WASM	Strojově čitelnější, a proto rychlejší, jazyk určující chování webové stránky
HTTP	Protokol popisující komunikaci mezi počítači
HTTPS	Šifrovaná nadstavba nad HTTP
Web API	Synonymum pro webovou službu
Autentizace	Ověření, že uživatel je tím za koho se vydává
Salt	Text zvyšující náhodnost hesla
Cookie	Data, která si stránka nebo API uložíla do prohlížeče
XSS	Útok vkládající utočníkův kód na důvěryhodnou webovou stránku
CSRF	Útok z nedůvěryhodné stránky využívající stavu přihlášen na důvěryhodné stránce
ACID	Žádoucí vlastnosti databáze (atomicita, konzistence, izolovanost, trvalost)
MVCC	Protokol pro paralelní práci s nějakým zdrojem založen na transakcích.

Část I

Teoretická část

Kapitola 1

Události a časové osy

V této práci se zabývám aplikací, která zobrazuje časové osy, proto vysvětlím, co taková časová osa vlastně je a co jsou události, které se na ní budou zobrazovat.

Časová osa je ve své podstatě linka, ve které každý bod reprezentuje nějaký okamžik, tedy nějaký den v roce v nějaký konkrétní čas. Na takové ose lze zaznamenat **události**, tedy nějaký popis dění v čase.

1.1 Druhy časových os

Časové osy lze dělit do kategorií a to například podle tvaru os (rovné, kruhové, spirálové, …), způsobu odsazení událostí (chronologicky, chronologicky relativně k nějaké události, s libovolným odsazením, logaritmicky, …), nebo počtu a posazení os (jedna osa, více os, jedna osa rozprostřená na více řádků, …). [1]

- *Kruhové* jsou v zásadě vhodné pro zaznamenání opakujících se událostí. Pěkným příkladem je Mayský kalendář.
- *Spirálové* se hodí pro esteticky příjemné zobrazení na malém prostoru. Obě varianty se spíše hodí pro ručně vytvořené sady událostí pro sdělení nějaké konkrétní informace. [1]
- Osy uspořádané *chronologicky relativně k nějaké události* lze použít pro porovnání dvou průběhů nějaké déle trvající události. Například pro srovnání dvou sportovních výkonů které se odehrály v různý čas. Tedy na dvou osách, které jsou zobrazeny nad sebou (stejně tak události „start“), je snadno vidět, v jakých úsecích se náskok získal, nebo ztratil. [1]
- Osy s událostmi *libovolně odsazenými* od sebe zobrazují pouze pořadí událostí a zanedbávají dobu, která mezi nimi uběhla. Naopak tím získá přehlednost v případech, kdy je na nějakých místech osy mnoho událostí blízko u sebe a na jiných prázdno. [1]
- *Rovné* osy uspořádané *chronologicky* jsou pravděpodobně právě ty, které si člověk představí pod pojmem časová osa. Vzdálenost mezi událostmi odpovídá množství času mezi událostmi a čas na zobrazovacím médiu plyne jen jedním směrem. [1]

1.2 Druhy časových událostí

Nejzákladnější typ události, označme ho jako *bodový*, reprezentuje dění v jeden konkrétní okamžik [2]. Nezajímá nás jak dlouho událost trvá, pouze kdy se udála. Lze ji tedy na ose reprezentovat jako bod.

Další základní typ, označme ho jako *trvající*, je charakteristický dobou trvání [2]. Například nějaká válka, nebo něčí vláda. Je důležité vědět, kdy událost začíná a kdy končí, protože pro bodové události může být relevantní, zda se udály souběžně s událostmi trvajícími. Například pokud byl člověk A zabit strážníkem B v době, kdy bylo vyhlášeno stanné právo, či nikoli. Takové události je vhodné zobrazovat jako úsečky. Dále máme trvající události, které pro potřeby časové osy začaly na počátku věků, nebo naopak stále trvají. Tyto události lze zobrazovat jako polopřímky.

Někdy může nastat problém, kdy nemáme úplné informace o události. Nemusí být znám přesný datum bodové události, ale pouze měsíc, či rok. Nebo není známo, kdy začala, nebo skončila trvající událost. Tyto události označme jako *nepřesné* a je vhodné je zobrazovat jako úsečky (či polopřímky), jejichž konce mohou být průhledné či jinak zbarvené, aby se indikovala tolerance nepřesnosti.

Kapitola 2

Rešerše podobných aplikací

Tato kapitola se zabývá souhrnem vlastností podobných aplikací nalezených během zpracování této práce.

2.1 TimelineSetter

TimelineSetter je nástroj, který z textového souboru, obsahujícího čas a popis události, vygeneruje zdrojový kód webové stránky, obsahující daná data reprezentovaná na časové ose [3].

Tato časová osa podporuje pouze bodové události, které zobrazuje jako úzké svislé pruhy uvnitř obdélníkové reprezentace osy. Po přesunutí kurzoru myši na pruh se zobrazí detaily o události. Jednotlivé události nejsou rozlišitelné bez zobrazení tohoto detailu.

2.2 Tiki-Toki Timeline Maker

Tiki-Toki Timeline Maker [4] má několik zobrazovacích módů. V této sekci se budu zabývat Category Band módem. Ten, jak je vidět zde 2.2, zobrazuje jednu osu v dolní části obrazovky, na které je posuvné okénko reprezentující rozsah zobrazených událostí. Události mají prostor ve zbytku obrazovky.

Každá událost je zobrazena jako rámeček s popiskem, který má na své spodní straně drobnou šipku ukazující na příslušný čas. Zdá se ale, že má problém s větším množstvím událostí blízko u sebe. Prostor pro události je rozdělen mezi jednotlivé kategorie událostí. Na časové ose v dolní části obrazovky se zobrazují barevné body reprezentující jednotlivé události pro jednoduchou navigaci na ose.

2.3 Timetoast timeline maker

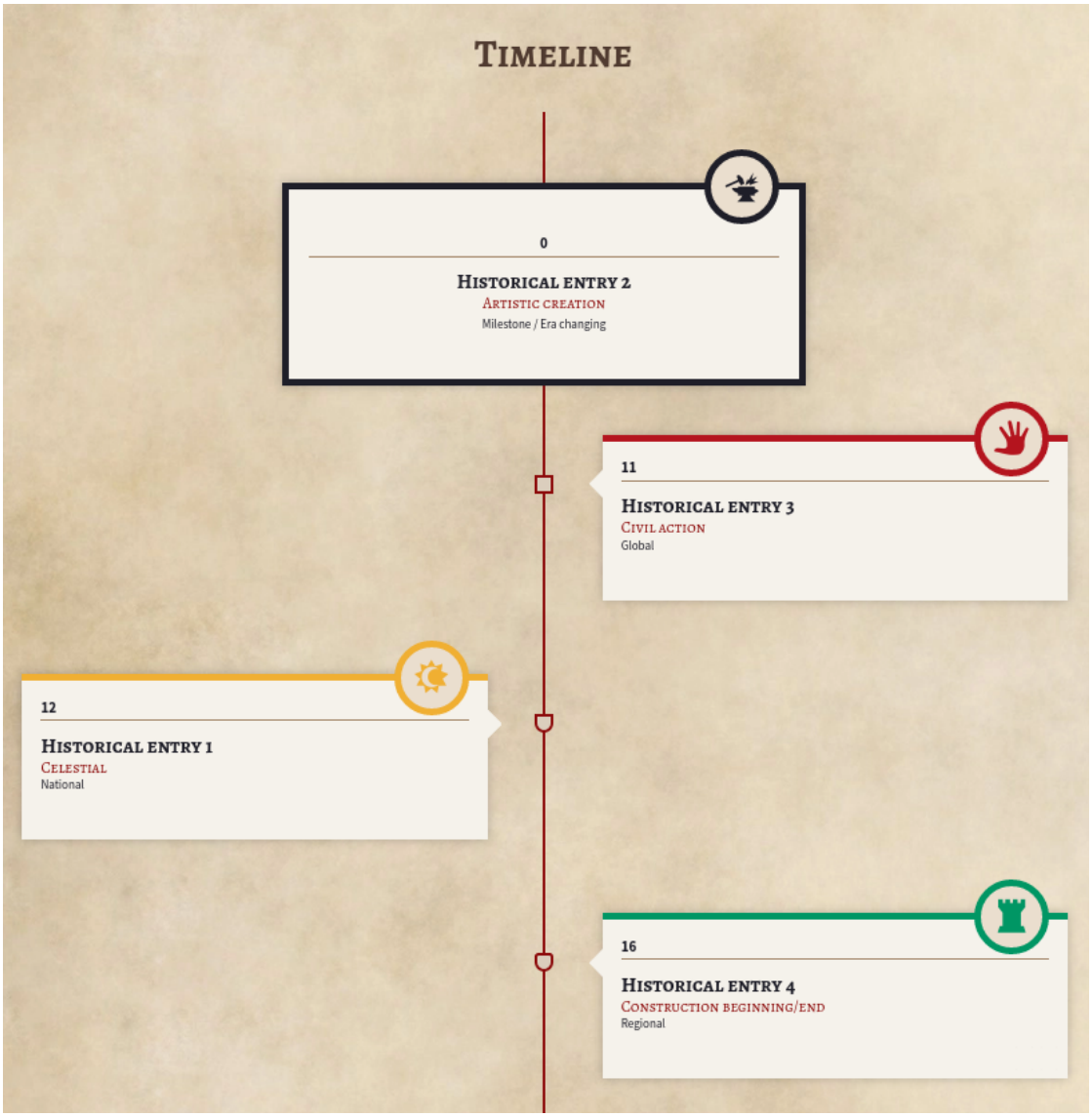
Timetoast timeline maker [5] stejně jako Tiki-Toki Timeline Maker [4] zobrazuje jednu osu v dolní části obrazovky. Ve zbylém prostoru se pak zobrazují jednotlivé události s rovnou čárou ukazující přímo dolů na konkrétní čas na ose. Viz 2.3.

Na rozdíl od Tiki-Toki Timeline Maker [4] se události rozmisťují nejdříve těsně nad osou a až při nedostatku místa se začnou přesunovat do řádků výše. Samotná osa si určí datum, před kterým, a datum, za kterým není zaznamenaná žádná událost. Mezi těmito daty pak leží posuvník, který má nastavitelný začátek i konec. Nad posuvníkem je zobrazený popisek dat časové osy, podle kterého se události zobrazují, jehož začátek a konec se řídí právě již zmíněným posuvníkem.

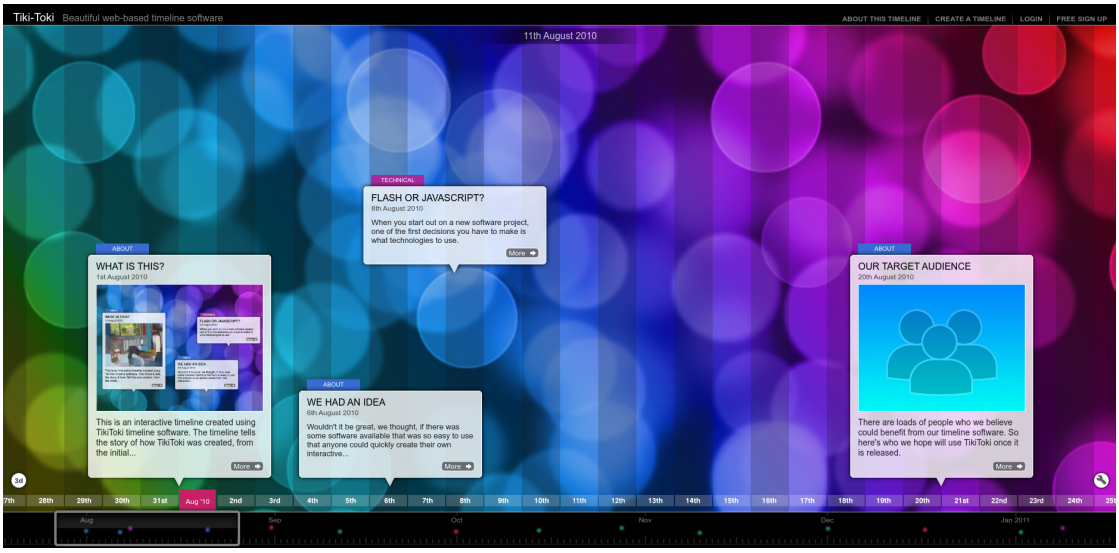
2.4 World Anvil

World Anvil [6] je nástroj, který umožňuje tvůrcům mít přehled o fiktivním světě. Tvůrce jednoduše začne tím, že si vybere co chce o světě napsat (popis postavy, státu, události, živočicha, ...). Nástroj pak tato data zobrazuje jako stránku principem podobnou Wikipedii [7].

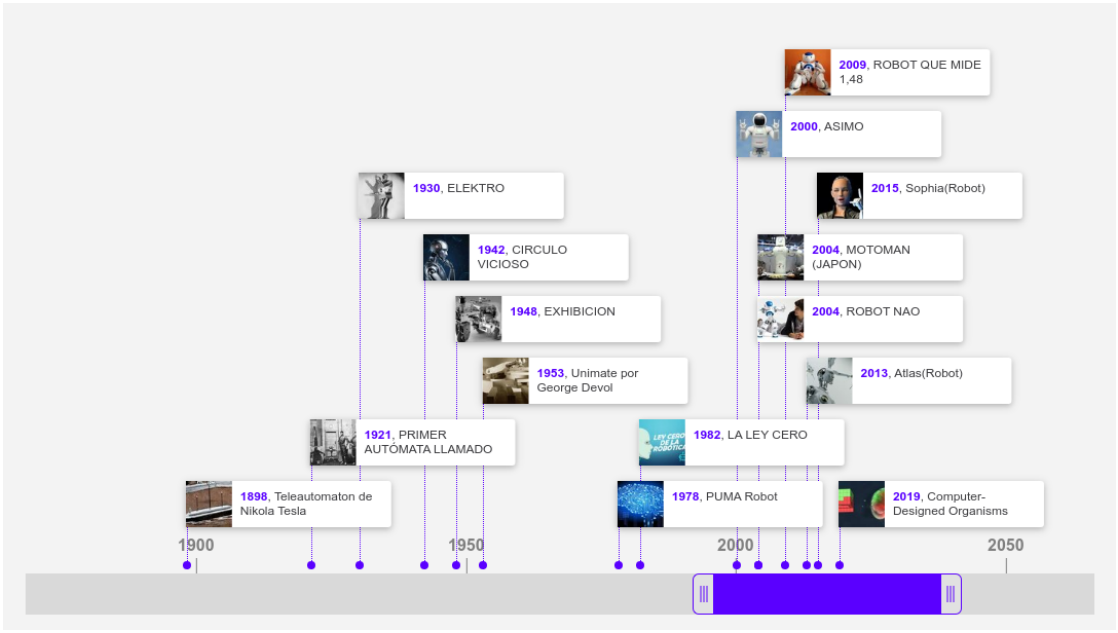
Nástroj také umožňuje tvorbu časových os 2.1. Tyto osy jsou svislé a události jsou zobrazeny z obou stran opět s malou šipkou směřující k ose. Události jsou na osách pouze seřazené a vzdálenosti mezi nimi neodpovídají délce času, který mezi nimi uplynul, ale jen výšce rámečku ve kterém je popisek události.



Obrázek 2.1 World anvil



Obrázek 2.2 Tiki-toki



Obrázek 2.3 Timetoast

Kapitola 3
Analýza podobných aplikací

TimelineSetter [3] se mi líbí snadným použitím, ale výsledná osa se mi zdá nepřehledná. Navíc se tento způsob řešení nehodí pro použití na webovém serveru.

World Anvil [6] má velmi pěkné rozhraní, jejich cíl je podobný jako cíl projektu, ale jejich časové osy jsou celkem nepřehledné pro větší množství událostí, protože každá událost zabírá velké množství místa na obrazovce a pravděpodobně z tohoto důvodu se rozhodli nepoužít chronologicky uspořádanou osu.

Tiki-Toki Timeline Maker [4] je již velmi podobný našemu cíli. Náznak množství událostí na ose v podobě teček na posuvníku velmi pomáhá přehlednosti na dlouhý osách, které nemáme zobrazené na obrazovce celé. Aplikace má ovšem problém se zobrazováním událostí co jsou velmi blízko u sebe, události místo přemístění prostě zobrazí přes sebe a to přehlednosti nepomáhá.

S tímto problémem si velmi dobře poradil Timetoast timeline maker [5], který události přesune o úroveň výše, aby se zamezilo překryvu. Z toho důvodu také používá tenkou svislou čáru, která označuje umístění události na ose.

Další zajímavý prvek této aplikace byl posuvník na časové ose. Úkol tohoto prvku byl posun osy, přiblížení osy a vizualizace, která část osy je zobrazená. Už tomu chyběla jen ta vizualizace množství událostí na ose, kterou předvedl Tiki-Toki Timeline Maker [4]

Kapitola 4
Webové technologie

Cílem projektu je vytvořit webovou aplikaci, tedy stránku. Webová stránka je text (kód), který říká internetovému prohlížeči, co a jak zobrazit a co a kdy dělat [8].

4.1 HTTP

Prohlížeč ovšem musí tento kód odněkud získat. Každá webová stránka je umístěna na nějakém počítači, který je dostupný přes internet [9]. Aby mohl prohlížeč získat kód webové stránky, musí nutně nějakým způsobem komunikovat s počítačem, který poskytuje veřejnosti onu stránku ke čtení.

Takové architektuře se říká klient-server. Cíl klienta je začít komunikovat se serverem a využívat nějakou službu. Server pak tuto službu poskytuje a odpovídá na klientovy dotazy. Klientem může být tedy počítač s webovým prohlížečem a server je počítač se spuštěnou aplikací poskytující webové stránky. [10]

Základní internetové protokoly umožňují posílat sled jedniček a nul [11]. Aplikace, které si tato data navzájem posílají, se musí dohodnout, co tato data vlastně znamenají. To není problém, pokud jsou například obě aplikace napsané stejným vývojářem. Pokud se ale bavíme o webových stránkách, pro které existuje mnoho různých prohlížečů a mnoho různých aplikací pro poskytování webových stránek, začne být problém podstatně komplikovanější.

Naštěstí existuje **HTTP** protokol, který přesně tuto problematiku řeší. HTTP tedy popisuje jak mezi sebou komunikují server a klient [12]. Posílané jedničky a nuly jsou v tomto případě chápány jako proud textu. Tento text, dále jen zpráva, je ve verzi HTTP/1.1 rozdělen na několik částí: metodu, cestu, verzi protokolu, stavový kód, hlavičky a tělo [13].

Dotaz a odpověď

Zprávy se pak dělí na dva typy, dotaz a odpověď. *Dotaz* 4.1 obsahuje metodu, cestu, verzi protokolu, hlavičky a volitelně tělo, přičemž součástí cesty mohou být další parametry. *Odpověď* 4.2 obsahuje verzi protokolu, stavový kód, stavovou zprávu, hlavičky a opět volitelně tělo. Pokud zpráva obsahuje tělo, jedna z hlaviček určuje, jaký způsobem se má tělo číst (v jakém je tělo formátu). [13]

Klient tedy pošle na server dotaz a server na tento dotaz odpoví. Server má nadefinováno, na jaké cestě pomocí jaké metody je jaká služba, a na základě dalších informací poskytnutých klientem v těle dotazu, v parametrech jako součást cesty nebo v hlavičkách se rozhodne, co bude dělat dál a co odpoví. Vzhledem k povaze protokolu je možné posílat mnoho typů neplatných dotazů, na které musí server nějak odpovědět. Z tohoto důvodu se v odpovědi posílají stavový

kód a stavová zpráva, které určují, zda byla odpověď poslána bez problémů, nastala interní chyba na serveru, dotaz byl špatně zformulován apod. [14]

Výpis kódu 4.1 Příklad dotazu

```
POST /test-api/category HTTP/1.1
Host: asterion-timelines.cz
Content-Type: application/json
Content-Length: 111

{   "name": "Category with icon",
    "description": "Loooong description of the category",
    "iconId": 1 }
```

Výpis kódu 4.2 Příklad odpovědi

```
HTTP/1.1 200 OK
Server: nginx/1.14.2
Date: Thu, 24 Jun 2021 21:17:05 GMT
Content-Type: application/json; charset=utf-8
Content-Length: 145
Connection: keep-alive
X-Powered-By: Express
Vary: Origin
Access-Control-Allow-Credentials: true
ETag: W/"91-GWIlZDJuxarzBvFOTvgTdGjTM0w"
Set-Cookie: connect.sid=abcdeXYZ; Path=/; HttpOnly; Secure

{   "id": 1,
    "name": "Místo",
    "description": "Libovolné místo, stát, ěmsto nebo kraj" }
```

HTTP a webové stránky

Prohlížeč obvykle při zadání běžné internetové adresy vyšle dotaz s metodou GET serveru, který se na této adrese nachází. To si lze jednoduše ověřit pomocí nástrojů na monitorování síťové komunikace. Server pak pošle odpověď se stavovým kódem označujícím, že je vše v pořádku (pokud je tedy vybraná stránka dostupná), a zdrojovým kódem webové stránky umístěným v těle odpovědi. Prohlížeč už pak díky hlavičkám a nějakým vlastním heuristikám pro rozpoznání textu ví, jestli má v odpovědi hledat webovou stránku nebo pdf soubor apod. [15]

Jelikož za odpovídání na dotazy je zodpovědná aplikace na straně serveru, často jsou stránky částečně sestaveny na serveru, odeslány a případně zapomenuty. Hodí se to například v případech, kdy jsou stránky šité na míru konkrétnímu uživateli, nebo pokud je pod mnoho adresami podobná stránka (např. internetová encyklopedie, multimediální přehrávač) apod. [16]

Další možností jak uživateli zobrazit unikátní stránku je vytvořit stránku, která získá potřebná data pomocí dalšího HTTP dotazu například pomocí technologie JavaScript. To se hodí především pro interaktivní aplikace, kdy není vhodné znovu načítat novou stránku při každém kliknutí na nějaké tlačítko. Případně lze použít kombinaci těchto dvou postupů. [16]

4.2 Webové stránky

Základní programovací jazyky

Nejjednodušší forma webových stránek si vystačí s HTML souborem [8]. Ten popisuje základní rozložení a prvky webové stránky (nadpisy, podnadpisy, odstavce, seznamy, tabulky, odkazy, obrázky, …) a nějaká metadata [17].

Pro větší kontrolu nad vzhledem zobrazených prvků vznikl jazyk CSS, který umožňuje nastavení nejrůznějších vlastností u každého prvku zvlášť nebo po skupinách (barva popředí a pozadí, rozměry, průhlednost, pozice, rozložení, změna vzhledu po přesunu kurzoru, …). [18]

Někdy ale ani to nestačí a je třeba vyžadované chování naprogramovat ručně. V takovém případě je zde JavaScript, plnohodnotný programovací jazyk, který umožňuje vkládat, mazat a manipulovat s prvky popsanými v HTML a jejich vlastnostmi specifikovanými v CSS. [19]

Dnes se již objevují další specifikace, které nahrazují JavaScript v místech, kde je velmi důležitá výpočetní rychlost. Například WebGL, rozhraní pro práci s OpenGL a přímou komunikaci s grafickou kartou [20], a WASM, jazyk navržený jako cíl kompilace z jiných jazyků a snadno a hlavně rychle pochopitelný prohlížečem a případně jinými aplikacemi [21].

Preprocesory

Jelikož se musí mnoho prohlížečů shodnout na podobě výše zmíněných technologií, trvá pochopitelně adopce lepších technologií celkem dlouho. Navíc ne každý se shodne na směru, kterým se webové technologie ubírají. Proto existují nástroje umožňující psát webové stránky v jiném jazyce (nebo existují přímo specializované jazyky) a následně je přeložit (zkompilovat) do HTML, CSS, JS a WASM. Těmto nástrojům říkáme preprocesory nebo kompilátory. [22] [23] [24]

Jedním z nich je například Sass. Tento nástroj kompiluje mimo jiné SCSS soubory (obdoba CSS s funkcionalitami navíc) do obyčejného CSS. [25] Dalším je TypeScript. To je nadmnožina JS která se do JS zkompiluje. [26]

Aby se nemusel každý nástroj spouštět zvlášť při každé změně kódu, existují nástroje, kterým se pouze nakonfiguruje, co se má kam a čím zkompilovat. Jeden z nejuniverzálnějších nástrojů pro tento účel je například Webpack. [27]

Kapitola 5

Technologické řešení webové vizualizace

Je několik základních způsobů jak něco zobrazit na webové stránce. Jedním je použít základní elementy HTML, nastylovat je pomocí CSS a nastavit jim chování pomocí JS. Druhou možností je použít jazyk SVG, který je velmi podobný HTML ovšem popisuje objekty vektorové grafiky se kterými lze pak manipulovat například pomocí CSS a JS. Další možností je použít softwarové vykreslování pomocí HTML elementu canvas (a jeho 2D kontextu). Neposlední možností je vykreslovat pomocí WebGL (HTML elementu canvas a jeho WebGL/3D kontextu). [28]

První varianta byla navržena k zobrazování statických prvků a její použití pro neustále se měnící strukturu je velmi pomalé [29]. *Druhá varianta* obsahuje sice jednodušší prvky pro manipulaci, ale animace těchto prvků stále spoléha na JS rozhraní DOM, které je pro větší počet prvků pomalejší [30]. *Třetí varianta* je na tom už o něco lépe. Je podstatně jednodušší vykreslování základních tvarů na libovolnou pozici ve vymezeném prostoru a je výrazně rychlejší než varianta první [29]. *Čtvrtá varianta* je sice výrazně rychlejší [29], ale má celkem složité rozhraní, jelkož oproti třetí variantě zpřístupňuje rozhraní grafické karty OpenGL.

Naštěstí existují knihovny, které zjednodušují práci s WebGL. Mnoho z nich je zaměřeno na 3d zobrazování nebo na tvorbu her (*three.js* [31], *Babylon.js* [32], *TWGL: A Tiny WebGL helper Library* [33], *vtk.js* [34], *PlayCanvas* [35], *Phaser - HTML5 Game Framework* [36]), se kterými by sice šlo zobrazovat pouze 2d objekty, ale knihovny jsou pak zbytečně velké a práce s nimi je pak zbytečně složitější. Pro zobrazování časových os bude stačit 2d vykreslování, které bude ovšem potenciálně zatížené velkým množstvím událostí potřebných k vykreslení. Proto je pro naše účely vhodná knihovna *PixiJS* [37], která má relativně jednoduché rozhraní a umí využívat WebGL kontext.

Kapitola 6

Autentizace uživatelů na webu

Tato kapitola se zabývá způsoby, jak bezpečně přihlásit uživatele. To mimo jiné znamená, že nikdo nesmí být schopen získat přístup k uživatelskému účtu bez znalosti přihlašovacího jména a hesla a my jako majitelé databáze nesmíme být schopni získat žádné heslo.

6.1 Identifikátor a heslo

Je mnoho způsobů jak provést autentizaci uživatele. Jedním ze způsobů je využít služby třetí osoby [38], jiným je například vygenerovat kód, který se odešle pomocí sms a který se po vyzvání zadá na příslušné stránce. Tato práce se ale bude zabývat autentizací pomocí unikátního identifikátoru a hesla, případně komunikací přes elektronickou poštu. Pro přihlášení tedy uživatel zadá svůj identifikátor a heslo do formuláře a stiskne tlačítko přihlásit. Úkolem je rozpoznat (ne)správnost hesla asociovaného se zadaným identifikátorem. Zároveň je třeba poskytnout možnost pro uživatele přihlásit se pouze pomocí HTTP rozhraní.

6.2 HTTPS

Pokud bychom nešifrovali komunikaci mezi uživatelem a serverem, tak by pro útočníka bylo velmi snadné identifikátor a heslo odposlechnout „po cestě“ pouze monitorováním síťového provozu. Proti odposlechu komunikace existuje technologie HTTPS [39]. Je to pouze nadstavba nad protokolem HTTP, která je schopná šifrované komunikace [39]. Po dešifrování je ovšem identická s HTTP [39]. Tím, že zašifrujeme komunikaci mezi uživatelem a API, znemožníme útočníkům komunikaci, tedi i identifikátor a heslo, odposlechnout.

6.3 Hašování

O šifrování komunikace se tedy postará HTTPS. Pokud by se ale útočník dostal k obsahu naší databáze, mohl by jednoduše použít získané heslo a přihlásit se. Nejen to, pokud uživatel použil stejné heslo i jinde, stačí útočníkovi zkusit použít ukradené heslo na známých sociálních sítích, obchodech apod.

Abychom tomu zamezili je třeba heslo v naší databázi takzvaně zahašovat. Jde o proces, který je deterministický a jednosměrný. To znamená, že pokaždé když zahašujeme heslo H dostaneme stejný haš H' a že když známe haš H' nejsme schopni získat původní heslo H jinak než zkoušením hesel a ověřením výsledku. [40]

6.4 Útok hrubou silou

Zahašujeme-li tedy každé heslo předtím než ho vložíme do databáze, můžeme správnost hesla kontrolovat tak, že heslo, které pošle uživatel zahašujeme a výsledný haš porovnáme s tím, který máme v databázi.

Útočník disponující ukradenými daty z naší databáze nebude schopný jen tak použít získaná hesla. Může ovšem vyzkoušet všechna možná hesla[1] a nalézt shodu s některým ze získaných hašů. Teprve poté by mohl nalezené heslo použít. Takovému útoku se říká útok hrubou silou. Pokud ovšem útočník získá data z naší databáze může takový útok provést offline a rychle, je proto vhodné použít takový algoritmus, který zabere hodně času. [40]

6.5 Slovníkový útok

Naším cílem je tedy donutit útočníka aby použil útok hrubou silou a aby mu takový útok trval co nejdéle. Útočník si ale může velmi efektivně pomoci. Slovníkový útok spoléhá na to, že uživatel si nastavil lidsky čitelné heslo, které se alespoň částečně sestává ze slov nalézajících se v jeho rodném nebo anglickém jazyce[2] nebo z již známých prolomených hesel. Útočník pak použije software na prolamování hesel, který je naprogramovaný, aby upravoval slova ze slovníku tak, jak by to udělal průměrný člověk. Tím se výrazně zvýší poměr mezi nalezenými hesly a počtem pokusů. [43]

Ať už způsob přihlášení zabezpečíme jakkoliv, útočník nakonec vždy může hádat hesla, která si uživatel nastavil. Proto je důležité vzdělávat uživatele o bezpečných heslech. Nejjistější metoda je hesla generovat strojově a použít správce hesel pro jejich zapamatování. Tím se do jisté míry zamezí útokům spoléhajícím na nespolehlivý generátor náhodných kombinací v lidské mysli.

6.6 Duhové tabulky

Další problém jsou duhové tabulky. Duhové tabulky jsou ve zkratce dlouhé seznamy hesel a jejich hašů. Ty jsou samy o sobě velmi dlouhé, ale princip duhových tabulek používá postup, který jejich velikost značně snižuje. Pro naše účely stačí tyto tabulky chápat jako šikovně indexovaný seznam zahašovaných hesel. Duhové tabulky spoléhají na jednoduchost uživatelských hesel a na známé postupy hašování. Každá duhová tabulka funguje jen pro nějaký konkrétní hašovací algoritmus a pro nějaké konkrétní rozmezí uživatelských hesel. Toto rozmezí může ovšem opět využít všech možných taktik, které jsou popsány výše (slovníky, předvídatelné úpravy). Jelikož jsou tyto tabulky i přes různé techniky velmi velké, jedna z nejlepších obran proti nim je použít dlouhé a komplikované heslo. [44]

6.7 Opakující se hesla

Uživatelům nic nebrání mít shodná hesla mezi sebou. Naopak, kdyby tomu něco bránilo, byl by to významný bezpečnostní problém. Pokaždé, když by někdo zadal heslo, které již někdo použil, tak by dotyčný měl informaci o tom, že někdo z přihlášených uživatelů takové heslo již má a to ještě předtím, než by získal obsah naší databáze. Musíme tedy počítat s tím, že různí uživatelé mají stejné heslo. To by se projevilo tak, že by příslušní uživatelé měli i stejný haš. Toho opět může útočník využít, jelikož tím, že zjistí heslo jednoho z nich, zjistí heslo všech. [45]

[1] Aby bylo nutné vyzkoušet opravdu všechny možnosti, je třeba použít kryptografický haš, který je pro takové použití vytvořený. Jako příklad uvedu *bcrypt*[41] a *argon2*[42].

[2] Mnoho stránek doporučuje při nastavení hesla použít velká písmena, čísla, speciální znaky apod. To sice pomáhá zpomalit útok hrubou silou, ale pokud útočník použije slovníkový útok, je velice předvídatelné kam uživatel vložil velké písmeno (na začátek), číslo (na konec, pravděpodobně 1-3 číslice) a speciální znaky (mezi slova a pravděpodobně _, &, ,).

6.8 Sůl (Salt)

S předchozími dvěma problémy nám může pomoci takzvaná sůl. Sůl je označení pro náhodný text, který se nějakým dobře definovaným způsobem přiloží k heslu předtím, než se zahašuje. Sůl nám pomůže zajistit složitost textu, který hašujeme i přesto, že zadané heslo bylo relativně jednoduché[3]. Je totiž vhodné použít sůl s dostatečnou délkou a složitostí[4]. Pokud navíc bude sůl unikátní a náhodná[5] pro každého uživatele, nebude možné identifikovat, kteří dva mají stejné heslo. Je ovšem důležité aby nebyla použita jedna stejná sůl napříč všemi uživateli. Bylo by pak možné vytvořit vlastní duhovou tabulku speciálně pro naší databázi a navíc by se nezamezilo ani problému s opakujícími se hesly. [45]

6.9 Identifikátor relace

Abychom nemuseli při každé operaci znovu a znovu (a složitě, jak jsme si již řekli) kontrolovat uživatelské jméno a heslo, nemluvě o jeho neustálém posílání, je vhodné využít k tomu náhodně vygenerovaný text (podstatně delší než heslo a vygenerovaný kryptograficky bezpečným generátorem), který po přihlášení uživateli pošleme zpět. Při každé další komunikaci již pro autentizaci bude stačit tento text (identifikátor relace). Tento identifikátor není nikde v databázi, protože je pouze v pracovní paměti webového rozhraní. Tím, že se uživatel odhlásí, smaže se identifikátor relace k němu příslušící a pro další komunikaci by se uživatel musel opět přihlásit, čímž by získal nový identifikátor. [46]

Jelikož je tento identifikátor náhodně generovaný, není jiná možnost než použít útok hrubou silou, který zároveň nebude benefitovat z taktik zaměřených na lidský faktor. A jelikož je velmi dlouhý je nerealistické, že by někdo uhodl něčí identifikátor v nějaké rozumné době. [46]

6.10 Sušenky (Cookies)

Je zřejmé, že přestože jsme autentizaci během většiny operací zjednodušili na poslání jednoho řádku textu, je nicméně stále potřeba tento řádek opakovaně odesílat. Pro zjednodušení této a podobných operací existují takzvané HTTP Cookies. Cookies je označení pro data která si aplikace komunikující přes HTTP a HTTPS pamatuje a je schopná je automaticky posílat jako součást všech dotazů. Jedním ze způsobů jak aplikace data získá je, že nějaká odpověď webového rozhraní obsahuje hlavičku, která říká aplikaci, jaká data si má zapamatovat. Příklad takové hlavičky můžeme vidět v kódu 6.1. [47] [48]

Výpis kódu 6.1 Příklad HTTP hlavičky nastavující cookie

```
Set-Cookie: session_id=some_random_number;
```

6.11 Šifrování cookies

Pokud by někdo zjistil, jaký uživatel používá identifikátor relace, mohl by se za něj po určitou dobu vydávat. To teoreticky lze útokem hrubou silou, ale je opravdu velmi nízká pravděpodobnost, že by se to někomu cíleně povedlo. Další možností je cookies odposlechnout. Tomu ovšem zabráníme tím, že používáme HTTPS protokol. Abychom si byli jisti, že používáme HTTPS protokol, je

[3] Stále je třeba aby měl uživatel složité heslo kvůli přímému slovníkovému útoku.

[4] Tím se značně omezí efektivita duhových tabulek jelikož je velmi obtížné předpovědět hesla do tabulky, která obsahují náhodné prvky. Bylo by třeba vytvořit duhovou tabulku šitou na míru pro každý seznam solených hašů. Ovšem i tak je složitost takového útoku obrovská, jelikož taková tabulka by musela být N krát větší pro zachování stejné efektivity (ne-li horší jelikož vyhledávání ve větší tabulce by trvalo déle), kde N je počet unikátních solí

[5] Tato sůl se při registraci uloží nešifrovaně k danému uživateli či heslu.

vhodné použít atribut `Secure`,jak je znázorněno v kódu 6.2, který zajistí, že pokud se uživatel připojuje přes protokol HTTP, cookie se neodešle [48].

6.12 Cross-site scripting

Cross-site scripting, dále jen XSS, je typ útoku, který vloží útočníkův kód do důvěryhodné webové stránky. To je možné díky špatně ošetřené komunikaci mezi uživatelem a webovým rozhraním. Nejdříve je třeba si uvědomit, jak může útočník vložit svůj obsah na stránku, kterou nevlastní. To lze velmi jednoduše. Například při vyhledávání na webu se zadá nějaký hledaný výraz do okénka a stiskne se tlačítko. Toto tlačítko pak přesměruje uživatele na jinou webovou adresu, která mimo jiné obsahuje i hledaný výraz. Tento výraz se pravděpodobně na stránce někde zobrazí. Pokud tedy útočník vytvoří takový odkaz, pak může svůj obsah vložit například právě do této části. Nebezpečí spočívá v tom, co do této části vloží. Pokud by tento hypotetický vyhledávač nebyl proti tomuto útoku odolný a útočník vložil do hledaného výrazu nějaký kód a hledaný výraz se pak vložil do stránky, webový prohlížeč by pak tento kus kódu jednoduše spustil. Z tohoto důvodu musí webové rozhraní ošetřit každý vstup od uživatele a přepsat ten vstup tak, aby prohlížeč kód nespustil. [49]

6.13 Cookies a javascript

Cookies jsou přístupné i přes rozhraní v jazyce javascript. To dává možnost útočníkům použít nějaký XSS a získat tak jednotlivé cookies. Jedním ze způsobů je ošetřit webové rozhraní před XSS. Dalším je nastavit dané cookie atribut `HttpOnly`, jak je znázorněno v kódu 6.2, který řekne prohlížeči aby nikdy nedovolil čtení této cookie přes JavaScript. 6.1. [47] [48]

Výpis kódu 6.2 Set-Cookie hlavička s atributy Secure a HttpOnly

```
Set-Cookie: session_id=some_random_number; Secure; HttpOnly
```

6.14 Posílání cookies jenom na naše rozhraní

Jak je popsáno výše, cookies jsou pouze data, která je prohlížeč, případně jiný HTTP klient, schopen posílat automaticky s každým dotazem na webové rozhraní. Pokud by ovšem nijak nerozlišoval mezi různými rozhraními, posílal by všechny cookies na všechna webová rozhraní. To v případě identifikátoru relace určitě nechceme. Naštěstí jsou cookies nastaveny tak aby se odesílali pouze na stejnou doménu, která danou cookie nastavila. Toto chování lze upravit atributy `Domain` a `Path`. 6.1. [47] [48]

6.15 Cross-site request forgery (CSRF)

CSRF je velmi zákeřný typ útoku. Útočník by, v případě úspěchu, byl schopen posílat dotazy na webové rozhraní pod jménem oběti. Představme si, že jsme vše zabezpečili proti výše uvedeným útokům. Uživatel se skutečně musí přihlásit, aby mohl s rozhraním dále pracovat, a cookies se neposílají na domény, které nám nepatří. Útočník skutečně nezíská naše přístupové údaje. On je totiž nepotřebuje. Když se přihlásíme na důvěryhodnou stránku, získáme tím cookie s naším identifikátorem. Pokud chceme pracovat s rozhraním, musíme mu spolu s HTTP dotazem poslat i cookie. Útočník tedy nepotřebuje znát žádné naše údaje, pokud je nějakým způsobem schopen zařídit, aby *z nějakého* počítače byl na rozhraní odeslán dotaz s útočníkem zadanými parametry a *s patřičou cookie oběti*. Jak je již uvedeno, patřičnou cookie má k dispozici pouze oběť a webové

rozhraní. Tedy pro útočníka je nejlogičtějším krokem donutit HTTP klienta oběti odeslat jím sestavený dotaz na příslušné rozhraní. To je ovšem bez patřičné ochrany velmi jednoduché. [49]

Útočník může například vytvořit vlastní webovou stránku. Každá webová stránka má možnost odeslat dotaz na libovolné rozhraní. Útočník na stránku umístil kód, který například odešle dotaz na webové rozhraní nějaké známé banky, který odešle nějaké množství peněz sobě na účet. Jelikož je dotaz směřovaný na rozhraní banky, tak pokud si prohlížeč pamatuje cookies poslané z rozhraní této konkrétní banky, odešle tyto cookies na rozhraní, které jednoduše uživatele ověří a odbaví dotaz. [49]

6.16 SameSite atribut

Kvůli CSRF vznikl atribut `SameSite`, který prohlížeči říká, jestli má patřičnou cookie odesílat pokud se uživatel nachází na jiné doméně, než ze které cookie pochází. atribut může nabývat několika hodnot[6] (`Lax`, `Strict`, `None`). Pokud je použita hodnota `None` cookie je zranitelná vůči CSRF. Pokud ovšem zmíněná cookie není schopná způsobit žádnou škodu, může to být žádoucí chování. Hodnota `Lax` zamezí odeslání cookie pokud se uživatel nachází na stránce třetí strany. Hodnota `Strict` funguje stejně jako hodnota Hodnota `Lax`, ale navíc ještě blokuje cookie, pokud by odeslání cookie bylo způsobeno kliknutím na odkaz na stránce třetí strany, přestože odkaz vedl na naše rozhraní. Tímto by se mělo minimalizovat riziko CSRF.

Tím že atribut nastavíme jako `Strict`, může to negativně ovlivnit uživatelovu zkušenost s naší aplikací. Představte si, že na nějaké stránce kliknete na odkaz na facebook a facebook se rozhodne, že vás kvůli tomu nepřihlásí. Aby se toto nestávalo je vhodné neautentizovat uživatele hned při vstupu na stránku, ale použít následné dotazy na rozhraní až z té konkrétní stránky.

Další doporučení říká, že je vhodné při návrhu rozhraní nepoužívat typu `GET` pro potenciálně nebezpečné dotazy, jelikož `GET` dotazy se standardně dají vložit do obyčejného hypertextového dotazu a je zde možnost, že při otevření odkazu z jiné aplikace (třeba poštovního klienta) v prohlížeči nebude správně detekováno a dojde k úspěšnému CSRF útoku i přes ochranu, kterou poskytuje `SameSite=Strict` atribut. [50]

[6] Je vhodné zmínit, že při absenci atributu `SameSite` při nastavení cookie, se většina moderních prohlížečů zachová tak, jako kdyby byla nastavena hodnota `Lax`

Kapitola 7

Množina časových dat světa Asterion

Jak jsem již zmínil, cílem tohoto projektu je vizualizace jednotlivých událostí na časové osy. Událostí může být libovolný popis dění. Jelikož je třeba události zobrazit na časové ose, je třeba je označit nějakým časovým údajem popisujícím dobu, kdy se událost odehrála. Většina událostí je popsána tak, že je lze označit jedním časovým údajem. Některé události ovšem chápeme jako nějaké období (například války), u kterých by bylo vhodné určit jak jejich začátek tak jejich konec. Ostatní události lze dekomponovat na ty již zmíněné. Například pokud bych chtěl popsat že období vlády P. Panovníka III. se dělí na několik částí (třeba podle věku nebo životních zkušeností), bylo by možné tuto skutečnost popsat separátními událostmi následovně:[1]

- Vláda P. Panovníka III. od časové značky 0 do časové značky 500
- Vláda P. P. III. v zastoupení S. Správce Hanebného od časové značky 0 do časové značky 100
- Vláda P. P. III. ve formě diktatury od značky 100 do značky 250
- Panovník se pohřešuje od 250 do 300
- Vláda P. P. III. ve formě parlamentní monarchie od 300 do 500

V předchozím příkladu používám časové značky namísto dat. Je to z toho důvodu, že svět Asterion má vlastní kalendář. Naštěstí je mnohem jednodušší než Gregoriánský kalendář. Rok (370 dní) na Asterionu se skládá ze 12 měsíců, kde každý má 30 dnů, a 10 speciálních svátkových dnů. Svátkové dny dělají kalendář lehce složitější. Svátkový den totiž nepatří k žádnému měsíci a tak nelze namapovat všechny dny v roce ke konkrétnímu datu ve formátu měsíc a den. Pro příklad uvedu data 30. Zelenec (poslední den, 5. měsíc) a 1. Ploden (první den, 6. měsíc), mezi kterými leží dva svátkové dny (Svátek letních duchů). Pro počítačové zpracování je tedy vhodnější zaznamenat pořadí dne v celém roce. Tedy 30. Zelenec je 152. den v roce a 1. Ploden je 155. den v roce. Při použití této metody je pak třeba ve frontendu tento datový typ převést na lidsky čitelné datum. Vzhledem k potřebě takového převodu, je jen logické, zahrnout do stejného čísla i rok a to tak, že časová značka bude reprezentovat počet dní od 1. dne roku 0 v Asterionu. Díky statickému počtu dní v roce je pak velice snadné spočítat rok i den v něm.

[1] Rád bych poukázal na možnost dekomponovat každou událost se značkami od a do na dvě samostatné události. (značka 0: Vláda P. P. III. začala; značka 500: Vláda P. P. III. skončila)

Kapitola 8

Rešerše a analýza databázových technologií

Cílem práce je mimo jiné navrhnout a implementovat databázi vhodnou pro množinu historických dat. Tato kapitola se zabývá rešerší databázových technologií a jejich případy užití.

8.1 Druhy databází

Existuje velké množství různých typů databázových technologií. Každá z nich umožňuje nějakým způsobem přidávat, číst, mazat a upravovat uchovaná data. Ty nejznámější z nich jsou hierarchické databáze, relační databáze, dokumentové databáze, asociativní databáze a grafové databáze. [51]

Hierarchické databáze

Hierarchické databáze byly původně vyvíjeny v 60. letech. Data jsou v nich uspořádaná do stromu, podobně jako v adresářové struktuře. Objekty v takových databázích mohou mít maximálně jednoho rodiče a libovolné množství potomků. Toto omezené uspořádání umožňuje rychlý přístup k objektům, ovšem značně omezuje schopnost databáze uchovávat provázaná data. [51]

Relační databáze

Relační databáze ukládají informace v tabulkách s předem definovanými sloupci. Nad těmito tabulkami se můžeme různě dotazovat, obvykle pomocí jazyka SQL. Tyto dotazy umožňují filtrovat záznamy na základě nějaké podmínky a kombinovat více tabulek do jedné například na základě shody záznamů v nějakém sloupci. [51]

NoSQL databáze

Moderní databáze, které nespadají do kategorie relačních databází, jsou často nazývány NoSQL databáze. To ovšem nutně neznamená, že daná databáze nepodporuje dotazovací jazyk SQL. Takové databáze uchovávají například JSON dokumenty, páry klíč-hodnota a grafy. [51]

Dokumentové databáze

Dokumentové databáze místo tabulek s řádky ukládají kolekce s dokumenty. Dokument je pak nějaký záznam neodlišný od souboru, který obvykle reprezentuje nějakou strukturu s hodnotami a klíči. Dokumenty tedy mohou být uloženy například ve formátech JSON, BSON, XML a podobně. K dokumentu lze pak přistoupit přes kolekci a unikátní identifikátor. Databáze by měla umožňovat přidávat, číst, mazat a upravovat dokumenty. Také by měla například umožňovat vyhledávání na základě jednotlivých hodnot v dokumentu. Velkou výhodou těchto databází je možnost nespecifikovat jednotnou strukturu dokumentů v kolekci. Populární dokumentovou databází jsou například MongoDB, Firebase a CouchDB [52]

Asociativní databáze

Asociativní databáze ukládají páry klíč-hodnota. Hodnoty jsou čteny, vkládány, mazány a upravovány na základě jejich unikátních klíčů. Hodnoty mohou být řetězce textů, čísla, jiné primitivní typy nebo dokonce složitější objekty jako seznamy, pole a haš tabulky. [53]

Grafové databáze

Grafové databáze obsahují dva typy entit. Uzly a hrany. Každý uzel může obsahovat nějaké vlastnosti, nebo nějaký dokument, dále mohou být kategorizovány a na základě toho případně omezeny nějakým schématem. Hrany obvykle obsahují odkaz na uzly ze kterého/kterých do kterého/kterých vedou, dále mohou obdobně jako uzly obsahovat nějaké vlastnosti, nebo dokumenty a mohou být nějakým způsobem kategorizovány a omezeny. nejpopulárnější grafové databáze používají způsob uložení těchto dat takový, že získání informací o sousedovi přes nějakou hranu je $O(1)$ zanedbáme-li velikost záznamu ohledně jednoho uzlu. [54]

Grafové databáze excelují v procházení uzlů přes hrany na základě nějakého rozhodovacího algoritmu. Pokud budeme v každém uzlu vědět přes kterou hranu budeme chtít přejít dále je složitost tohoto algoritmu $O(k)$, kde k je počet průchodů. Naproti tomu pokud bychom to samé chtěli udělat v relační databázi, je třeba pro každý průchod vyhledat souseda v tabulce. Pokud tedy máme indexovanou tabulku uzlů pomocí nějakého vyhledávacího stromu v relační databázi je složitost takového průchodu $O(k \log n)$, kde n je počet uzlů v tabulce. Díky tomu také excelují v případech, kdy přesně nevíme, přes kterou hranu chceme projít příště. Tedy při hledání cest splňujících nějaké požadavky, či hledání nejkratších cest mezi uzly. [54] [55]

Pozorování

Při čtení o různých typech databází je vidět, že zatím není přesně definováno, kde přesně leží hranice mezi jednotlivými typy. Například hierarchické databáze lze velmi snadno simulovat například pomocí dokumentových databází díky kolekcím. Stejně tak by šla hierarchická databáze simulovat pomocí asociativní databáze a to tak, že pro každou hodnotu v hierarchické databázi by byla klíčem do asociativní databáze její cesta. Tyto „simulace“ lze dokonce provést bez ztráty efektivity databáze. Tato „šedá zóna“ dokonce způsobuje nejasnost mezi zdroji jestli je Windows Registry hierarchická nebo asociativní databáze [53] [51].

Dále lze například simulovat grafovou databázi pomocí relační databáze, kde každá návštěva souseda by byla provedena sloučením tabulek a filtrací. Tento převod by ovšem velmi zhoršil efektivitu zejména při přecházení přes vícero hran. Důvod proč to zmiňuji je ovšem ten, že přestože by došlo k výraznému zhoršení efektivity v případech, ve kterých „nativně“ grafová databáze exceluje, došlo ve veřejné diskuzi k rozhodnutí, že grafovou databázi neurčuje implementace ale, podle definice, její rozhraní [55]. Nadstavba relační databáze s rozhraním grafové databáze by tedy podle definice měla být obojím.

MongoDB zároveň tvrdí, že jejich dokumentová databáze je schopná pracovat jako mnoho jiných typů jako jsou asociativní, relační, objektové, grafové a geoprostorové [52].

Těmito příklady chci ukázat jak nejasná je hranice mezi různými databázemi. Z toho co jsem viděl jsem přesvědčen o tom, že je zbytečné slovíčkařit nad jednotlivými definicemi typů databází, ale že je při hodnocení efektivity databáze důležité to, jaký problém je třeba v dané doméně řešit, tím pádem jaké dotazy budete databázi pokládat a přes jaké rozhraní, a jaký způsob uložení těchto dat bude nejvhodnější pro takové použití daného rozhraní.

8.2 Analýza typů databázových technologií

Množinu časových dat světa Asterion lze v zásadě chápat jako graf, kde vrcholy jsou jacísi aktéři a události a kde hrany spojují dohromady skupiny aktérů a aktéry s událostmi. Ovšem rozhodl jsem se použít relační databázi ze dvou důvodů. Jsem s prací s relační databází obeznámen lépe než s prací s jinými typy. Vzhledem k návrhu aplikace, kde jsme rozhodli, že se budou zobrazovat pouze data, která jsou „velmi blízko“ nějakému vrcholu je relační databáze více než dostačující, jelikož dojde pouze k malému množství sloučení tabulek na dotaz.

8.3 Jednotlivé relační databáze

Tato sekce se zabývá zhodnocením vybraných relačních databází ze seznamu od `db-engines.com` [56]. Databáze byly vybrané na základě pořadí v seznamu a licence, tedy tři v době psaní této práce nejpopulárnější open-source databáze.

MySQL

MySQL používá klient-server architekturu. Vícevláknový serveru je díky systémovému kernelu možné snadno používat i na více procesorech. Zároveň server podporuje několik platforem mezi něž patří i Linux a Windows Server Operating Systems. Samozřejmostí jsou transakce splňující ACID principy. Oproti PostgreSQL dosahuje MySQL lepšího výkonu při čtení a má zároveň mnohem snadnější instalaci a použití a také větší komunitu. Nevýhodou může být horší výkon při vkládání většího počtu dat najednou a pomalé souběžné operace pro čtení a zápis. [57]

PostgreSQL

Stejně jako MySQL, PostgreSQL používá klient-server architekturu. PostgreSQL ovšem na rozdíl od MySQL používá pro každého připojeného uživatele separátní proces. Samozřejmostí jsou opět transakce splňující ACID principy. Oproti MySQL, PostgeSQL podporuje, aby jednotlivé pohledy měly vlastní keš, a tím urychluje práci s jednotlivými pohledy. Dále používá MVCC technologii pro zachování konzistence dat při paralelním zpracování, která je lepší a rychlejší než pouze používání zámků. MySQL také podporuje MVCC za použití InnoDB. PostgreSQL je považována za rychlejší databázi právě díky vyspělejší technologii paralelního zpracování a zároveň je užívána ve finančních aplikacích pro striktní plnění ACID principů. PostgreSQL také umožňuje načítání nativních dynamických knihoven implementujících nové funkce za běhu programu [58]. Oproti MySQL je ale zároveň mnohem více náročnější na paměť, jelikož každý separátní proces (tedy každé připojení) vyžaduje až 10MB. [57]

SQLite

SQLite je velmi odlišný od MySQL a PostgreSQL, oproti nim totiž používá pouze jediný soubor pro uchování všech potřebných struktur. Zároveň nevyžaduje žádnou instalaci a obvykle se používá jako knihovna vložená přímo do aplikace, která databázi využívá a každá taková aplikace si vytvoří vlastní soubor s danou databází. Dále je SQLite jedna z nejmenších databází. Je dokonce menší

než většina knihoven pro komunikaci s klient-server databázemi. Díky tomu je vhodná pro použití do zařízení s nízkou výpočetní silou a do jednotlivých aplikací jako úložiště lokálních dat. Hlavní nevýhodou je ovšem neschopnost paralelního přístupu a použití více uživateli. Pravděpodobně díky tomu také neexistuje mnoho poskytovatelů databáze SQLite jako cloud službu. [57]

8.4 Analýza relačních databází

Je zřejmé, že SQLite není příliš vhodná pro použití jako databázi pro http službu, převážně kvůli nedostatku paralelního přístupu. PostgreSQL a MySQL jsou mnohem lepší kandidáti, ovšem PostgreSQL je velmi složitá databáze k instalaci a údržbě. MySQL je dnes již plnohodnotnou databází, jejíž mnoho údajných nevýhod (například chybějící fulltextové vyhledávání [57]) již neplatí [59]. Z těchto důvodů a také proto, že je open-source, použiji MySQL databázi.

Kapitola 9

Rešerše a analýza webových serverů

Cílem práce je také navrhnout a implementovat webové rozhraní pro komunikaci s frontend aplikací. Tato kapitola se tedy zabývá rešerší jednotlivých webových serverů. Na základě seznamu nejpoužívanějších webových serverů [60] a jejich licencích jsem do rešerše uvažoval tyto aplikace: Apache HTTP Server a Nginx. Dále jsem uvažoval následující kombinace jazyků a knihoven pro implementaci HTTP rozhraní: Java, JRE a Spring Boot; JavaScript, Node.js a Express.js; C++ a POCO C++ libraries.

9.1 Apache HTTP Server

Apache HTTP Server byl nejpopulárnějším serverem od roku 1996. Díky tomu má výbornou dokumentaci a integeaci s jinými aplikacemi. Apache používá dynamicky načítaný systém modulů, které umožňují interpretaci různých jazyků bez použití externí aplikace. [61]

Apache, oproti Nginx, umožňuje lokální nastavení uvnitř složky s obsahem pomocí `.htaccess` souboru. Nginx nepodporuje žádný takový mechanismus a vyžaduje aby byl upraven centrální konfigurační soubor. To ovšem znamená, že Apache v tomto ohledu vyžaduje větší režii při každém dotazu. [61]

Apache byl navržen primárně pro použití jako server pro webové stránky, a proto je primárně uzpůsoben k mapování částí URI adres do míst v souborovém systému. Pro nasměrování do jiných lokací je třeba vytvořit různé aliasy. [61]

9.2 Nginx

Nginx získal popularitu díky lehčímu využití prostředků a možnosti jednoduchého růstu za použití minimálních nároků na hardware. Nginx exceluje v poskytování statického obsahu a je navržen pro přeposílání dotazů na dynamický obsah jinému softwaru, který se pro tento účel hodí více. [61]

Na rozdíl od Apache, Nginx byl od začátku navržen k poskytování statického obsahu a zároveň k fungování jako proxy server. Díky tomu je uzpůsoben primárně k překladu URI adres do jiných lokací a až sekundárně poskytuje možnost přesměrovat nějakou URI do souborového systému. Dále Nginx, oproti Apache, neposkytuje dyamický systém modulů a je třeba nově zkompilovat nginx spolu s potřebnými moduly. [61]

9.3 Node.js

Node.js používá událostmi řízený jedno-vláknový neblokující I/O model. To znamená, že umožňuje volat funkce, které pro své dokončení musejí čekat po nějakou dobu, tak aby volající takové funkce mohl spustit více takových funkcí najednou dříve než dojde k dokončení těch předchozích. Díky tomu je Node.js velmi rychlý s použitím jednoho vlákna a málo paměti. Node.js je napsaný a využívá jazyk JavaScript. [62]

9.4 Spring Boot

Spring Boot umožňuje rychle zprovoznit kvalitní aplikaci z větší části pomocí Java anotací a XML kofiguračních souborů. Díky tomu je vhodný pro aplikace, u kterých je třeba rychle vytvořit prototyp a očekává se růst aplikace. Spring Boot dále využívá více vláken najednou a tím tak dokáže využít procesor na maximum. [62]

9.5 POCO

POCO knihovny pro C++ obsahují všechny důležité moduly pro tvorbu výkonných webových serverů. Umožňují jednotný přístup k několika různým relačním databázím, obsahují klienty k MongoDB a Redis databázím, umožňují šifrování, multiplatformní přístup k souborovému systému a manipulaci s cestami, paralelní přístup, parsování XML a JSON dat a hlavně podporu HTTP, TCP a SSL. [63]

Ovšem stejně jako pro každý případ užití C++ platí, že C++ umožňuje vysoký výkon, ale vývoj takové aplikace trvá mnohem déle a je mnohem složitější. Z toho důvodu je dost dobře možné vytvořit méně výkonnou aplikaci než využít specializované a jednodušší nástroje.

9.6 Analýza

Při výběru této technologie pro tuto práci hrála velkou roli jednoduchost použití. POCO knihovny nejsou jednoduché na použití už z důvodu, že jsou napsané pro C++. Spring Boot se pyšní svou jednoduchostí, ale dle mé zkušenosti je třeba napsat mnohem více boilerplate kódu než v Node.js s knihovnou Express.js. Dále jsem mnohem více obeznámen s programováním v JavaScriptu než v Javě. Z těchto důvodů jsem vybral Node.js jako technologii pro implementaci webového rozhraní.

Jelikož ale máme k dispozici jen jeden server, na kterém je třeba mít frontend i backend, je třeba nějak rozdělit URI k jednotlivým aplikacím. Pro tento účel se hodí spíše Nginx než Apache, protože je vhodnější pro použití jako proxy. Navíc je možné nakonfigurovat SSL zabezpečení pouze pomocí Nginx a není pak nutné řešit šifrování pomocí Node.js.

Část II

Praktická část

Kapitola 10

Návrh prototypu

Aplikace se často dělí na backend a frontend. Zjednodušeně řečeno frontend je část aplikace, se kterou interaguje uživatel, a backend je zbytek. Cílem této práce je navrhnout a nasadit backend aplikace, proto se v této kapitole zaměřím na návrh databáze a webového API.

Popíši zde požadavky, případy užití a doménový model, navrhnu diagram aktivit, diagram tříd, databázový model a model nasazení.

10.1 Požadavky

- Autorizace uživatelů
- Přidávání, odebírání a editace událostí autorizovanými uživateli
- Přidávání, odebírání a editace štítků (tagů) autorizovanými uživateli
- Označení událostí jedním nebo vícero tagy
- Vyhledávání tagů
- Vyhledávání událostí podle zvoleného tagu nebo nějakého aliasu, který označuje jeden nebo více tagů
- Neautorizovaný uživatel nesmí být schopen manipulovat s daty v databázi.
- Nikdo nesmí být schopen z databáze nebo během autorizace získat něčí heslo.
- Nikdo nesmí být schopen změnit něčí heslo bez přístupu k tomuto účtu.
- Nikdo nesmí být schopen použít něčí účet bez znalosti přihlašovacích údajů.
- Snadno pochopitelná API

10.2 Případy užití

V této sekci se zaměřím na jednotlivé příklady použití navrhované API.

Vyhledání událostí podle tagů a metatagů (filtrů)

Jedním z hlavních cílů projektu je v určité podobě zobrazovat události týkající se konkrétních tagů a metatagů. (Tuto dvojici budu nadále nazývat filtry.) Ovšem pro účely úpravy dat v databázi by bylo vhodné implementovat i vyhledávání podle názvu a popisu události s případným časovým omezením.

Vyhledání filtrů

Aby bylo možné filtrovat události s konkrétním filtrem, je třeba takový filtr napřed vyhledat. Filtry by mělo být možné hledat podle jména a popisu. V nejlepším případě by se vyhledané filtry seřadily podle podobnosti s hledaným výrazem. Tyto filtry by měly být seskupeny do kategorií určujících zda se jedná o místo, osobu a podobně.

Zobrazování ikon

Součástí událostí, filtrů a kategorií může být i ikona. Odkaz na ni by měl být posílán jako součást vyžádaných dat.

Autorizace uživatele

Uživatel se bude moci do API zaregistrovat, přihlásit a odhlásit. Po přihlášení bude skrze API možné provádět operace, které jsou běžně nedostupné, předpokládaje, že k tomu byl uživatel oprávněn. Zároveň není po uživateli API vyžadováno autentizaci řešit vyjma přihlášení, odhlášení a zaregistrování, jelikož při každém volání nějaké API služby se pomocí cookies provede autentizace a následná autorizace automaticky.

Přidávání dat

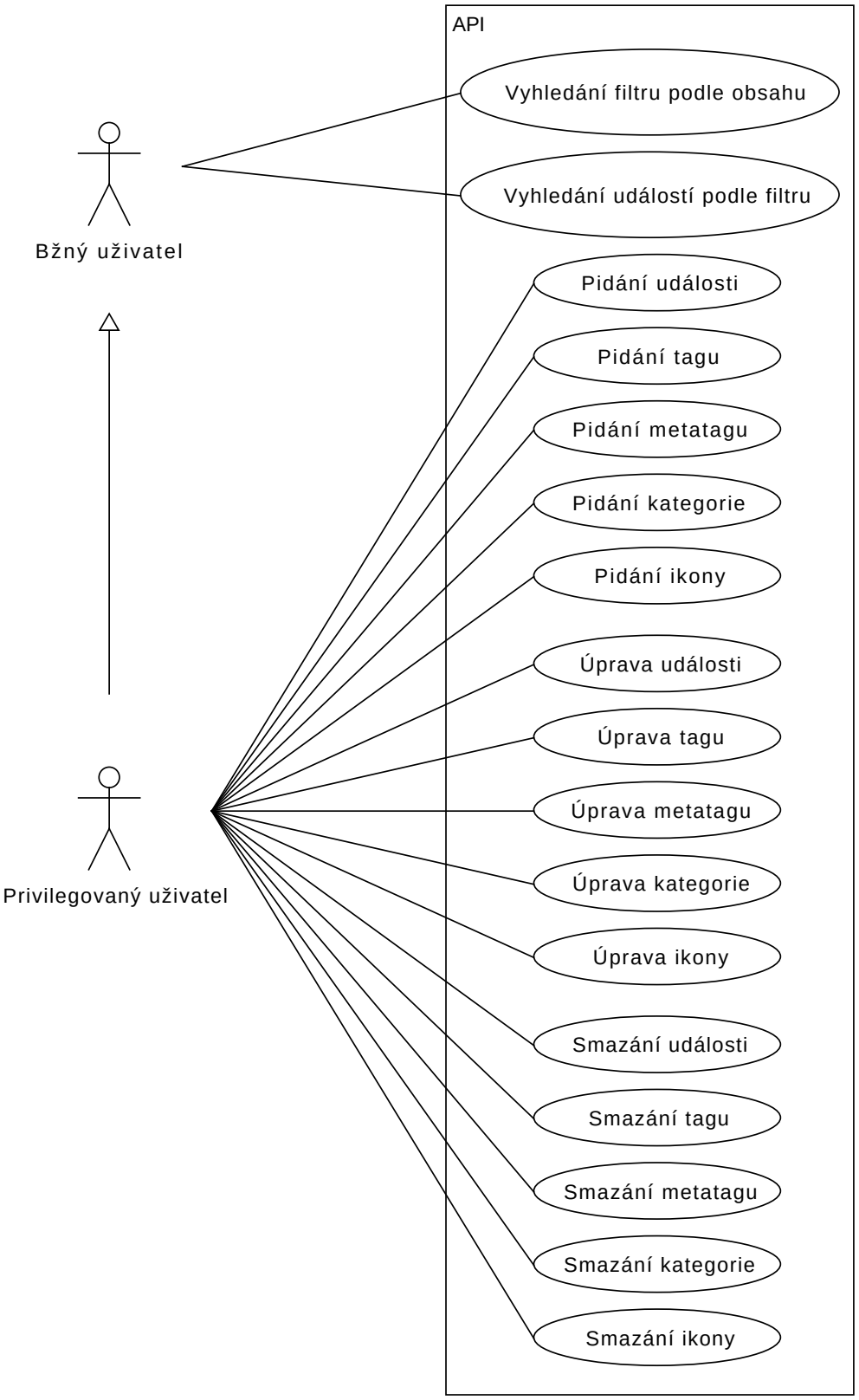
Privilegovaný uživatel by měl být schopen přidávat události, filtry, kategorie a ikony.

Úprava dat

Privilegovaný uživatel by měl být schopen upravovat události, filtry, kategorie a ikony.

Smazání dat

Privilegovaný uživatel by měl být schopen mazat události, filtry, kategorie a ikony.



Obrázek 10.1 UseCase diagram

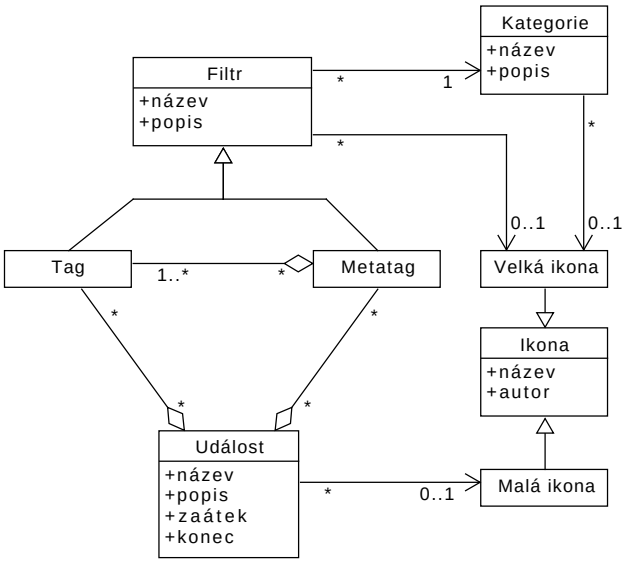
10.3 Doménový model

Historická data lze namodelovat následujícím způsobem. Mějme třídu Filtr označující subjekty zájmu (podstatná jméno ve větách) a třídu Událost označující jejich nebo jich se týkající činnosti (slovesa ve větách).

Subjekty zájmu lze ale dále seskupovat nebo přejmenovávat. Například božstvo může být označení pro skupinu více subjektů (jednotlivých bohů) a každý bůh může mít více jmen, pod kterými je znám. Rozdělme tedy třídu Filtr na podtřídy Tag a Metatag a propojme je m:n vazbou. Událost pak bude mít m:n vazbu jak s třídou Tag tak s třídou Metatag.

Dále lze jednotlivé subjekty kategorizovat do velkých skupin, jako například místa, osoby apod. Mějme tedy třídu Kategorie, která bude přiřazena každému filtru.

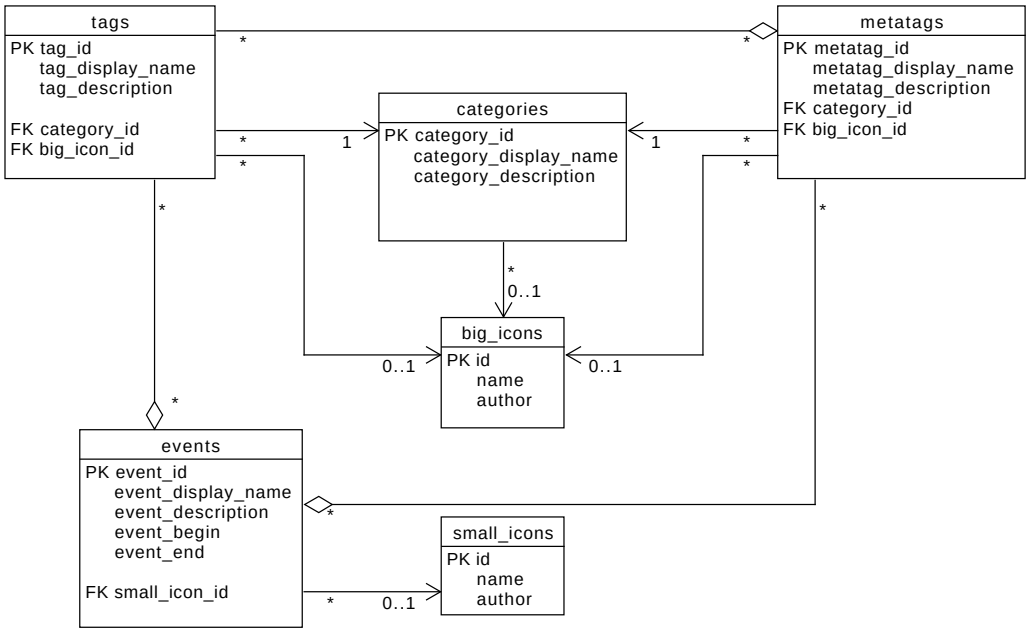
Vzhledem k tomu, že vytváříme grafickou aplikaci, je třeba jednotlivé filtry a události vizualizovat. Proto mějme třídu Ikona a její podtřídy Velká a Malá ikona. Velké ikony jsou přiřazeny k filtrům a kategoriím a malé ikony k událostem. Kategorie mají přiřazenou ikonu pro případ, že samotný filtr žádnou přiřazenou nemá a v případě, že ani kategorie nemá žádnou přiřazenou, bude vybrána jedna výchozí. Zároveň bude existovat i výchozí malá ikona pro případ, že událost nebude mít žádnou konkrétní přiřazenou.



Obrázek 10.2 Doménový model

10.4 Datový model

Na základě předchozího doménového modelu je navržen následující model dat, ovšem byla z něj vyjmuta dědičnost tříd, která se vyskytuje až v business logice. Z modelu je vynechána tabulka uživatelů obsahující id, uživatelská jména, emaily a haše hesel a tabulka povolení obsahující id. Tyto dvě tabulky mezi sebou mají m:n vazbu. Databázový model je pak pouze datový model s dekomponovanými m:n vazbami pomocí dekompoziční tabulky.



Obrázek 10.3 Datový model

10.5 Návrh webového rozhraní

Rozhraní by mělo obsahovat několik hlavních sekcí. Události, tagy, metatagy, filtry, kategorie a ikony jsou sekce zodpovědné za přidávání, úpravu a mazání dat. Filtry jsou ovšem pouze pomocná sekce pro vyhledávání mezi tagy a metatagy zároveň. Další sekce by měla být zodpovědná za autentizaci uživatelů.

Události

Tato sekce by měla být přístupná na adrese `/event` (tedy `url.cz/api/event`) a pomocí http metod GET, POST, PUT a DELETE zprostředkovává funkce vyhledávání pomocí id, přidávání, úpravu a mazání událostí. Dále by měla obsahovat dotazy na vyhledání pomocí id filtru a pomocí textového obsahu události.

Tagy a metatagy

Tyto sekce by měly být přístupné na adresách `/tag` a `/metatag` a pomocí http metod GET, POST, PUT a DELETE zprostředkovávat funkce vyhledávání pomocí id, přidávání, úpravu a mazání tagů a metatagů. Dále by měla obsahovat dotazy na vyhledání podle textového obsahu a

vyhledání asociovaných protějšků, tedy v sekci Tagy vyhledání všech metatagů obsahujících daný tag a v sekci Metatagy vyhledání všech tagů obsažených v daném metatagu.

Filtry

Tato sekce by měla být zpřístupněna na adrese `/filter` a pomocí http metody GET zprostředkovávat vyhledávání jak mezi tagy tak mezi metatagy pomocí id. Dále by měla obsahovat dotaz na vyhledání tagů a metatagů podle textového obsahu.

Kategorie

Tato sekce by měla být přístupná na adrese `/category` a pomocí http metod GET, POST, PUT a DELETE zprostředkovává funkce vyhledávání pomocí id, přidávání, úpravu a mazání kategorií tagů a metatagů. Dále by měla obsahovat dotazy na vyhledání pomocí textového obsahu kategorie a na vypsání všech kategorií.

Přihlašování

Tato sekce by měla být přístupná na adrese `/sign` a pomocí dotazů `/sign/up`, `/sign/in` a `/sign/out` zajišťovat registraci, přihlášení a odhlášení uživatele. Dále by měla obsahovat dotazy pro změnu informací přihlášeného uživatele, žádost o reset hesla a žádost o odstranění účtu a údajů s ním spojených.

Uživatelé

Tato sekce by měla být přístupná na adrese `/user` a pomocí http metod GET, POST, PUT a DELETE zprostředkovává funkce vyhledávání pomocí id, přidávání, úpravu a mazání uživatelů. Dále by měla obsahovat dotazy na přidávání a odebírání oprávnění uživatelům.

Kapitola 11
Implementace aplikace

Aplikace je napsána v jazyce JavaScript a spouštěna pomocí Node.js. Pro práci s http protokolem byla vybrána knihovna express.js, pro hašování hesel knihovna argon2, pro přihlašování uživatelů knihovna passport.js a pro komunikaci s databází knihovna mysql2.

11.1 Přihlašování uživatelů

Předtím než bylo rozhodnuto pro zřízení vlastní databáze uživatelů bylo v plánu použít autentizaci poskytovanou společností Google. To se ale ukázalo být překvapivě obtížné kvůli dokumentaci, která neobsahovala řešení žádného problému, který se naskytl, a všemožné návody (přímo od googlu) byly roztroušeny napříč různými sekcemi dokumentace google api schovanými mezi mnohými jinými návody na management skupin a jiných nám nepotřebných funkcí. Postup v návodu, na který se odkazovalo nejčastěji, vyžadoval použití knihovny, která byla v jiných částech dokumentace zastaralá, a na novou knihovnu se nepodařilo najít dokumentaci zabývající se základním použitím.

Rozhodlo se tedy pro passport.js. Dokumentace passport.js sice také nebyla perfektní, byla velmi stručná a některé věci si člověk musel domyslet, ale alespoň byla konzistentní. Passport.js navíc umožňuje jednoduché začlenění jiných přihlašovacích strategií do frameworku express.js v budoucnu.

Zatímco passport.js se stará o udržování sezení přihlášeného uživatele, při použití lokální strategie je třeba naimplementovat databázi uživatelů. Ta v našem případě obsahuje uživatelské jméno, email a haš hesla. Navíc se k ní váží práva uživatelů. Hašování a ověřování hesel je implementováno pomocí algoritmu argon2, který vyhrál Password Hashing Competition (PHC) v roce 2015 a mimo jiné automaticky solí hesla náhodným řetězcem.

11.2 Code injection

Aby nedošlo k žádnému útoku pomocí code injection, je používán systém připravených příkazů (prepared statements) v mysql, který odděleně pošle SQL příkaz a poté data. Databáze tak od sebe dokáže jednotlivé části odlišit a nedojde tak záměně uživatelských dat za příkaz. Zároveň je na celou API nasazen překladač, který nahrazuje podezřelé znaky za jejich html reprezentaci (html entitu), aby nedošlo k vložení kódu na stránku ze strany nevhodných dat vložených do databáze.

11.3 Vrstvy

API je rozdělena do tří hlavních vrstev. První, databázová, vrstva má na starost komunikaci s MySQL, druhá má na starost překlad dat z databáze do správného formátu a získání doplňujících dat pomocí databázové vrstvy a třetí vrstva zprostředkovává tato data na specifických umístěních v API pomocí express.js frameworku.

11.4 Finální podoba API

Implementovaná API obsahuje následující dotazy (získané pomocí dotazu `GET::/endpoints`) zapsané v JSON formátu 11.1 tak, že každý string představuje jeden dotaz. Metoda dotazu je rovna obsahu stringu. Cesta dotazu je rovna zápisu všech názvů předků pole (s názvem `"::"`) obsahujícího daný string. Pro příklad uvedu první dotaz. Dotaz má metodu POST a je přístupný na adrese `/sign/up`, respektive `http://www.asterion-timelines.cz/api/sign/up`.

Formát parametrů

Dotazy s metodou GET jsou parametrizovány pomocí URL encoded query a dotazy s jinou metodou jsou parametrizovány pomocí těla dotazu ve formátu JSON. Výjimkou jsou dotazy obsahující jako parametr soubor (především dotazy týkající se ikon) ty jsou parametrizovány pomocí těla dotazu ve formátu multipart form data.

Použité metody

Dotazy s metodou GET slouží k získání dat, s metodou POST k vytvoření nových dat, s metodou PUT k úpravě již existujících dat a s metodou DELETE ke smazání dat. Výjimkou mohou být dotazy na adresách obsahujících prefix `/sign` a `/ping`.

Výpis kódu 11.1 Finální podoba API

```
{   "/sign": {
        "/up": { "::": [ "POST" ]},
        "/in": { "::": [ "POST" ]},
        "/out": { "::": [ "POST" ]},
        "/logged": { "::": [ "GET" ]},
        "/user": { "::": [ "GET", "PUT" ]},
        "/check_login": { "::": [ "GET" ]},
        "/remove_account": { "::": [ "POST" ]}},
    "/user": { "::": [ "GET", "PUT", "POST", "DELETE" ],
        "/byLogin": { "::": [ "GET" ]}},
    "/permit": { "::": [ "GET" ],
        "/byName": { "::": [ "GET" ] },
        "/all": { "::": [ "GET" ] }},
    "/tag": { "::": [ "GET", "POST", "PUT", "DELETE" ],
        "/byContent": { "::": [ "GET" ]},
        "/metatags": { "::": [ "GET", "PUT" ]}},
    "/metatag": { "::": [ "GET", "POST", "PUT", "DELETE" ],
        "/byContent": { "::": [ "GET" ]},
        "/tags": { "::": [ "GET", "PUT" ]}},
    "/filter": { "::": [ "GET" ],
        "/byContent": { "::": [ "GET" ]}},
    "/category": { "::": [ "GET", "POST", "PUT", "DELETE" ],
        "/byContent": { "::": [ "GET" ]},
        "/all": { "::": [ "GET" ] }},
```

```
    "/event": { "::": [ "GET", "POST", "PUT", "DELETE" ],
        "/byContent": { "::": [ "GET" ] },
        "/byFilterId": { "::": [ "GET" ] }},
    "/icon": {
        "/tags": { "::": [ "GET", "POST", "PUT", "DELETE" ],
            "/neededBy": { "::": [ "GET" ] },
            "/byPath": { "::": [ "GET" ] },
            "/byName": { "::": [ "GET" ] },
            "/all": { "::": [ "GET" ] },
            "/default": { "::": [ "PUT" ] },
            "/default_from_path": { "::": [ "PUT" ] },
            "/default_from_id": { "::": [ "PUT" ] }},
        "/events": { "::": [ "GET", "POST", "PUT", "DELETE" ],
            "/neededBy": { "::": [ "GET" ] },
            "/byPath": { "::": [ "GET" ] },
            "/byName": { "::": [ "GET" ] },
            "/all": { "::": [ "GET" ] },
            "/default": { "::": [ "PUT" ] },
            "/default_from_path": { "::": [ "PUT" ] },
            "/default_from_id": { "::": [ "PUT" ] }}},
    "/ping": {
        "/api": { "::": [ "GET" ] },
        "/logged": { "::": [ "GET" ] },
        "/db": { "::": [ "GET" ] },
        "/log": { "::": [
            "GET", "POST", "PUT", "DELETE",
            "HEAD", "CONNECT", "OPTIONS",
            "TRACE", "PATCH" ]}},
    "/endpoints": { "::": [ "GET" ],
        "/permitted": { "::": [ "GET" ] }}}
```

Hlavní dotazy

Dotazy na adresách s prefixem `/permit`, `/tag`, `/metatag`, `/category`, `/event`, `/filter`, `/icon/tags` a `/icon/events` pojmenujme jako *hlavní dotazy* a tyto prefixy jejich adres chápejme dále také jako `/prefix`. Hlavní dotazy pak dodržují následující strukturu.

Vyžadované parametry - GET

Hlavní dotazy s metodou GET na adresách `/prefix` vyžadují parametr `id` a v odpovědi vrátí JSON s daty daného objektu. Pro vyhledávání mohou existovat dotazy s metodou GET na adresách `/prefix/bySomething`, kde `something` vyjadřuje parametr, pomocí kterého vyhledáváme (všimněte si malého písmene na začátku parametru). Tyto *vyhledávací* dotazy mohou vracet pole (stále ve formátu JSON) pokud se vyhledává pomocí parametru, který není unikátní. Dotazy na adresách `/prefix/all` pak vrací všechna známá data tohoto typu.

Vyžadované parametry - PUT

Hlavní dotazy s metodou PUT jsou parametrizovány parametry téměř totožnými s těmi, které získáme dotazem `/prefix` s metodou GET. Parametr `id` je vždy povinný, tím se určí, který objekt chceme upravit. Ostatní parametry jsou nepovinné (mohou být `undefined`).

Rozdíl je ve vnořených objektech v dotazech s formátem JSON. Pokud získaný objekt z odpovědi obsahuje jiný objekt nebo pole objektů (11.2), dotaz s metodou PUT použitý k úpravě

takového objektu požaduje pouze jeho id nebo pole všech id (11.3) (všiměme si změny z `"object"` na `"objectId"` a změny z `objects` na `objectIds`).

Při úpravě zmíněných polí je třeba vložit pole všech id, které chceme vložit, API pak vytvoří dvě nová pole (co vložit a co smazat), která využije při odbavení dotazu jako změnové vektory. Pokud tedy byla na serveru uložená data 11.2 a odeslali jsme dotaz s metodou PUT a daty 11.3 budou data v databázi vypadat následovně 11.4. Pokud chceme pole zachovat nezměněné nebudeme parametr `objectIds` vůbec posílat a nebo ho nastavíme jako `undefined`.

Výpis kódu 11.2 Odpověď na dotaz GET s id 5

```
{
   "id": 5,
   ...,
   "object": {
      "id": 7,
      ...
   },
   "objects": [
      {
         "id": 4,
         ...
      }
      {
         "id": 2,
         ...
      }
   ]
}
```

Výpis kódu 11.3 Tělo dotazu s metodou PUT

```
{
   "id":5,
   ...,
   "objectId": 89,
   "objectIds": [1, 2, 8]
}
```

Výpis kódu 11.4 Odpověď na druhý dotaz GET s id 5

```
{
   "id": 5,
   ...,
   "object": {
      "id": 7,
      ...
   },
   "objects": [
      {
         "id": 4,
         ...
      }
      {
         "id": 2,
         ...
      }
   ]
}
```

Vyžadované parametry - POST

Hlavní dotazy s metodou POST jsou parametrizovány stejně jako s metodou PUT ovšem parametru `id` je nepovinný a ignorován. Dále jsou pro tyto dotazy povinné následující parametry.

- `/category`
  - `name`
  - `description`
- `/tag, /metatag`
  - `name`
  - `description`
  - `categoryId`
- `/event`
  - `name`
  - `description`
  - `begin` (počet dnů od dne 0 roku 0, `{ return (370 * rok) + den }`)
- `/icon/tags, /icon/events`
  - `name` (multipart form data text)
  - `icon` (multipart form data soubor)

Vyžadované parametry - DELETE

Hlavní dotazy s metodou DELETE jsou parametrizovány pouze parametrem `id`, který je samozřejmě povinný.

Zvláštní případ - /filter

Všiměme si, že sekce `/filter` obsahuje pouze dotazy s metodou GET. Je to proto, že objekty typu filter jsou poze ke čtení a vznikají kombinací dat typu tag a metatag. **Pozor**. Id filtru není zpětně kompatibilní s id tagu a id metatagu. Pokud potřebujete id tagu nebo metatagu je nutné [1] použít vyhledávací dotazy tagů a metatagů.

Zvláštní případ - /icon/events a /icon/tags

Ikony mají několik dotazů, které neodpovídají výše uvedené struktuře. Dotaz `/neededBy` vypíše všechny události popřípadě tagy, metatagy a kategorie, které jsou závislé na ikoně specifikované pomocí parametru `id`. Dotaz `/default` nastaví výchozí ikonu na soubor odeslaný v parametru `icon`. Dotaz `/default_from_path` nastaví ikonu (stejného typu) nacházející se na adrese specifikované parametrem `path` jako výchozí ikonu. Dotaz `/default_from_id` nastaví ikonu (stejného typu) mající id specifikované parametrem `id` jako výchozí ikonu.

[1] Pokud není jiná cesta, lze id převést následujícím způsobem, ovšem tento způsob není součástí dokumentace a může se v budoucnu změnit. Id filtru je string. Pokud má na začátku podtržítko ('_'), je třeba ho odstranit. Tím získáme id metatagu. Pokud nemá podtržítko, získali jsme id tagu. Je ovšem nutné dodat, že id tagů a metatagů jsou čísla.

Zvláštní případ - /tag/metatags a /metatag/tags

Tyto dotazy slouží k úpravě m:n vztahu mezi tagy a metatagy a dodržují konvenci ohledně vnořených polí a objektů zmíněnou výše. Pro upřesnění uvedu, že dotazy `/tag/metatags` pracují s parametrem `/metatagIds` a dotazy `/metatag/tags` pracují s parametrem `/tagIds`.

Endpoints

V sekci přístupné na adrese `/endpoints` jsou dva dotazy. Dotaz `/endpoints::GET` poskytuje seznam všech dotazů stejně jak je popsáno výše a `/endpoints/permitted::GET` poskytuje objekt reprezentující slovník mezi dotazem a hodnotou true/false označující zda má uživatel povolení využívat tento dotaz.

Ping

Sekce na adrese `/ping` jsou následující dotazy. Dotazy `/ping/api::GET` a `/ping/db::GET` odpoví statusem 200 a zprávou Ping. Pokud ne tak není v provozu api, respektive není v provozu spojení api s databází. `/ping/logged::GET` odpoví statusem 200 a zprávou Ping pokud je uživatel přihlášen, pokud ne tak status odpovědi bude 401. Dotazy `/ping/log` zapíší data dotazu do logu podezřelých. Tato skupina dotazů vznikla kvůli objeveným voláním adresy `/jsonws/invoke` z neznámé Ruské IP a je zamýšlena jako místo kam přesměrovat podezřelé dotazy pomocí web serveru nginx.

Sign

Sekce na adrese `/sign` slouží primárně k manipulaci se stavem přihlášení a manipulaci s daty přihlášeného uživatele.

Přihlašování, odhlašování a registrace

Registrace, přihlašování a odhlašování jsou zprostředkovány pomocí dotazů `/sign/up`, `/sign/in` a `/sign/out`. Všechny dotazy používají metodu POST. Registrace vyžaduje parametry `username`, `email` a `password`. Přihlášení vyžaduje parametry `login`, který může představovat jak username tak email, a `password`. Odhlášení nevyžaduje žádné paraetry.

/sign/check_login

Dotaz přijímá parametr `login`, který představuje email nebo uživatelské jméno. Pokud je email, nebo uživatelské jméno již obsazené, vrátí odpověď `{free: false}` a v opačném případě `{free: true}`. Momentálně přihlášený uživatel nemá s dotazem souvislost. Dotaz byl vytvořen za účelem zjednodušení implementace frontendu, konkrétně registračního formuláře.

/sign/remove_account

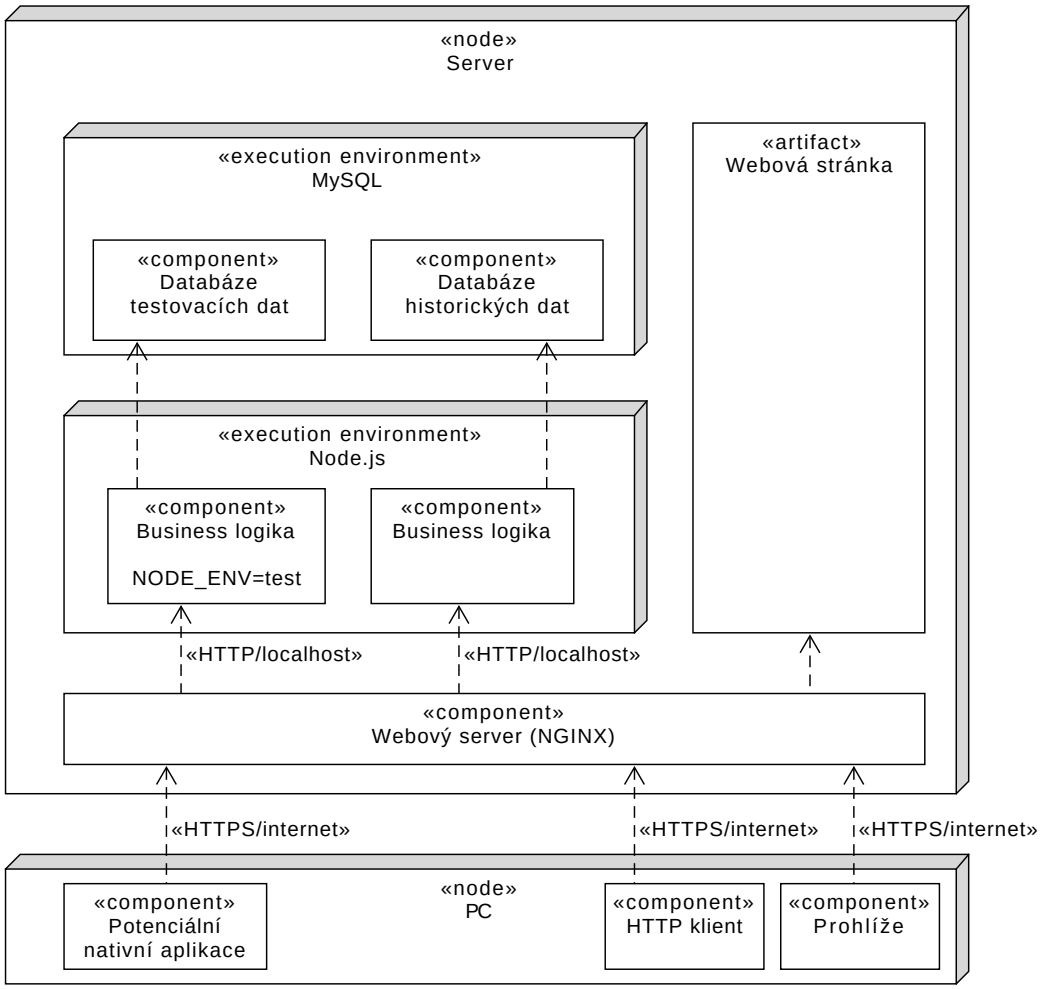
Dotaz odstraní účet přihlášeného uživatele, tedy jeho email, uživatelské jméno, haš jeho hesla a informaci o tom, jaká vlastnil oprávnění. **Frontend aplikace by měly být velmi obezřetné s použitím této operce.**

/sign/user

Dotaz `/sign/user` pomocí metody GET slouží k získání informací o přihlášeném uživateli. Dotaz byl vytvořen za účelem zjednodušení implementace frontendu, konkrétně indikaci přihlášeného uživatele. Dotaz `/sign/user` pomocí metody PUT slouží k úpravě informací o přihlášeném uživateli (uživatelské jméno, email, heslo) a pracuje s parametry `username`, `email` a `password`, pokud je parametr vynechán zůstane informace, kterou představuje nezměněna.

11.5 Nasazení aplikace

Aplikace je nasazena na zadavatelem poskytnutém serveru. Zabezpečenou komunikaci přes https s uživatelem zprostředkovává webový server nginx[64], který na adrese `/` vrací statickou webovou stránku (zkompilovaný frontend) a na adrese `/api/` přeposílá dotazy Node.js aplikaci (backend), která následně využívá MySQL databázi. Druhá dvojice Node.js aplikace a MySQL databáze je pak přístupná na adrese `/test-api/` pro účely testování.



Obrázek 11.1 Model nasazení

11.6 Testování aplikace

Testování aplikace probíhalo v několika krocích. Po implementaci nějakého dotazu se funkce ručně otestovala v aplikaci Postman. Poté byl přidán nový dotaz do sekvence automatizovaných testů (opět v aplikaci Postman), která byla spouštěna pravidelně během práce na projektu.

Byla také vytvořena webová stránka, na které se testovalo praktické použití API a odolnost vůči code injection útokům.

Dále byla otestována slabost proti útokům popsaným v kapitole 6 týkajících se technologii cookies. Cookies nejsou přístupné přes javascript, cookies se neposílají na jiné webové stránky než a cross-site request forgery se nezdařilo.

Pomocí nástroje MySQL Workbench bylo zkontrolováno, že hesla jsou zahašovaná algoritmem argon2 a uživatelé se stejným heslem mají odlišný haš.

Jsem tedy přesvědčen, že API je zabezpečena.

Závěr

Cílem této práce bylo vytvořit technické zázemí pro projekt časových os světa Asterion a databázi událostí a webové API, čili backend. Byla provedena rešerše konkurence a byly nalezeny velké nedostatky v jednotlivých podobných aplikacích. Webové rozhraní bylo implementováno, zpřístupněno veřejnosti, otestováno a shledáno funkčním a bezpečným.

Ukázalo se, že přestože jsem byl více obeznámen s jazykem JavaScript, bylo by mnohem snadnější naučit se TypeScript kvůli typovému systému, který by zamezil častým chybám v průběhu vývoje způsobených voláním funkcí s nesprávným datovým typem, především `null`, `undefined`, `NaN` apod. Dále by stálo za uvážení použít kombinaci grafové a dokumentové databáze pro ukládání událostí souvisejících s aktéry a případně i jinými událostmi spolu s dokumentem ke každému aktéru a události s detailnějším popisem. Ovšem návrh takové aplikace a datového modelu by byl zcela odlišný.

Aplikaci by nemělo být příliš složité rozšířit o encyklopedické znalosti o jednotlivých událostech, tazích a metatazích, které by posléze šlo zobrazovat po kliknutí na danou událost, tag či metatag, například pomocí přidání sloupce do databáze s odkazem na separátní wiki stránky. Dále je tato aplikace pro svět Asterion specifická pouze svým kalendářem. Databáze ukládá časová razítka, která lze interpretovat libovolným způsobem, a je tedy možné po drobné úpravě webového rozhraní ukládat například historická data reálného světa.

Bibliografie

1. BREHMER, M.; LEE, B.; BACH, B.; RICHE, N. H.; MUNZNER, T. Timelines Revisited: A Design Space and Considerations for Expressive Storytelling. *IEEE Transactions on Visualization and Computer Graphics*. 2017, roč. 23, č. 9, s. 2151–2164. Dostupné z DOI: 10.1109/TVCG.2016.2614803.
2. FULDA, J.; BREHMER, M.; MUNZNER, T. TimeLineCurator: Interactive Authoring of Visual Timelines from Unstructured Text. *IEEE Transactions on Visualization and Computer Graphics*. 2016, roč. 22, č. 1, s. 300–309. Dostupné z DOI: 10.1109/TVCG.2015.2467531.
3. PRO PUBLICA INC. *TimelineSetter* [online]. 2011 [cit. 2021-03-22]. Dostupné z: http://propublica.github.io/timeline-setter/.
4. WEBALON LTD. *Tiki-Toki Timeline Maker* [online]. 2021 [cit. 2021-03-23]. Dostupné z: https://www.tiki-toki.com.
5. TIMETOAST TIMELINES. *Timetoast timeline maker* [online]. 2021 [cit. 2021-03-23]. Dostupné z: https://www.timetoast.com.
6. THE WORLD ANVIL TEAM. *World Anvil* [online]. 2020 [cit. 2021-03-27]. Dostupné z: https://www.worldanvil.com.
7. WIKIMEDIA FOUNDATION INC. *Wikipedia: The free encyclopedia* [online]. 2004 [cit. 2021-03-27]. Dostupné z: https://www.wikipedia.org.
8. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *What is the difference between webpage, website, web server, and search engine?* [Online]. 2021 [cit. 2021-11-02]. Dostupné z: https://developer.mozilla.org/en-US/docs/Learn/Common_questions/Pages_sites_servers_and_search_engines.
9. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *What is a web server?* [Online]. 2021 [cit. 2021-11-02]. Dostupné z: https://developer.mozilla.org/en-US/docs/Learn/Common_questions/What_is_a_web_server.
10. BRITANNICA, THE EDITORS OF ENCYCLOPAEDIA. *client-server architecture* [online]. 2021 [cit. 2021-11-10]. Dostupné z: https://www.britannica.com/technology/client-server-architecture.
11. FRYSTYK, Henrik. *The Internet Protocol Stack* [online]. 1994 [cit. 2021-11-10]. Dostupné z: https://www.w3.org/People/Frystyk/thesis/TcpIp.html.
12. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *HTTP* [online]. 2021 [cit. 2021-11-02]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/HTTP.
13. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *HTTP Messages* [online]. 2021 [cit. 2021-11-10]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/HTTP/Messages.

14. THE INTERNET SOCIETY. *Hypertext Transfer Protocol – HTTP/1.1* [online]. 1999 [cit. 2021-11-10]. Dostupné z: https://www.w3.org/Protocols/rfc2616/rfc2616.html.
15. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *How the Web works* [online]. 2021 [cit. 2021-11-02]. Dostupné z: https://developer.mozilla.org/en-US/docs/Learn/Getting_started_with_the_web/How_the_Web_works.
16. OMNISCI, INC. *Server-Side Rendering* [online]. 2021 [cit. 2021-11-02]. Dostupné z: https://www.omnisci.com/technical-glossary/server-side-rendering.
17. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *HTML: HyperText Markup Language* [online]. 2021 [cit. 2021-11-02]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/HTML.
18. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *CSS: Cascading Style Sheets* [online]. 2021 [cit. 2021-11-02]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/CSS.
19. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *About JavaScript* [online]. 2021 [cit. 2021-11-02]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/JavaScript/About_JavaScript.
20. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *WebGL: 2D and 3D graphics for the web* [online]. 2021 [cit. 2021-12-13]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/API/WebGL_API.
21. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *WebAssembly* [online]. 2021 [cit. 2021-12-13]. Dostupné z: https://developer.mozilla.org/en-US/docs/WebAssembly.
22. KOLCE, James. *10 Languages That Compile to JavaScript* [online]. 2018 [cit. 2021-12-13]. Dostupné z: https://www.sitepoint.com/10-languages-compile-javascript/.
23. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *CSS preprocessor* [online]. 2021 [cit. 2021-12-13]. Dostupné z: https://developer.mozilla.org/en-US/docs/Glossary/CSS_preprocessor.
24. AGEEK.DEV. *What Languages Support WebAssembly?* [Online]. 2021 [cit. 2021-12-13]. Dostupné z: https://ageek.dev/wasm-langs.
25. THE SASS TEAM, AND NUMEROUS CONTRIBUTORS. *CSS with superpowers* [online]. 2021 [cit. 2021-12-13]. Dostupné z: https://sass-lang.com.
26. MICROSOFT. *TypeScript is JavaScript with syntax for types.* [Online]. 2021 [cit. 2021-12-13]. Dostupné z: https://www.typescriptlang.org.
27. DEVELOPERDRIVE STAFF. *6 BEST BUILD TOOLS FOR FRONTEND DEVELOPMENT* [online]. 2021 [cit. 2021-12-13]. Dostupné z: https://www.developerdrive.com/best-build-tools-frontend-development/.
28. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *Graphics on the Web* [online]. 2021 [cit. 2021-12-13]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/Guide/Graphics.
29. WALTERS, Abraham. *Canvas vs. DOM vs. WebGL vs. ImpactJS vs. PixiJS Particle Test* [online]. 2013 [cit. 2021-04-18]. Dostupné z: https://github.com/quidmonkey/particle_test.
30. SMUS, Boris. *Performance of canvas versus SVG* [online]. 2009 [cit. 2021-12-13]. Dostupné z: https://smus.com/canvas-vs-svg-performance/.
31. THREE.JS AUTHORS. *three.js* [online]. 2021 [cit. 2021-04-03]. Dostupné z: https://threejs.org.
32. CATUHE, David. *Babylon.js* [online] [cit. 2021-04-03]. Dostupné z: https://www.babylonjs.com.

33. TAVARES, Gregg. *TWGL: A Tiny WebGL helper Library* [online]. 2019 [cit. 2021-04-03]. Dostupné z: https://twgljs.org.
34. KITWARE INC. *vtk.js* [online]. 2021 [cit. 2021-04-03]. Dostupné z: https://kitware.github.io/vtk-js/.
35. PLAYCANVAS LTD. *PlayCanvas* [online]. 2021 [cit. 2021-04-03]. Dostupné z: https://playcanvas.com.
36. PHOTON STORM LTD. *Phaser - HTML5 Game Framework* [online]. 2021 [cit. 2021-04-03]. Dostupné z: https://phaser.io.
37. GOODBOY DIGITAL LTD. *PixiJS* [online]. 2021 [cit. 2021-04-03]. Dostupné z: https://www.pixijs.com.
38. JONES, William. *An Introduction to Authentication as a Service (AaaS)* [online]. 2019 [cit. 2021-12-13]. Dostupné z: https://medium.com/@williamjonescodes/introduction-to-authentication-as-a-service-aaas-11e70f8c83c5.
39. MISHRA, Abhishek. *How HTTP and HTTPS Work?* [Online]. 2020 [cit. 2021-12-13]. Dostupné z: https://medium.com/swlh/how-http-and-https-work-4c689e1ea369.
40. CHEATSHEETS SERIES TEAM. *Password Storage Cheat Sheet* [online]. 2021 [cit. 2021-12-13]. Dostupné z: https://cheatsheetseries.owasp.org/cheatsheets/Password_Storage_Cheat_Sheet.html.
41. PROVOS, Niels; MAZIERES, David. A Future-Adaptable Password Scheme. *Proceedings of 1999 USENIX Annual Technical Conference* [online]. 1999, s. 81–92 [cit. 2021-04-22]. Dostupné z: https://archive.org/details/1999-proceedings-freenix-track-atc-monterey/page/81/mode/2up.
42. AUMASSON, Jean-Philippe. *The Password Hashing Competition* [online]. 2019 [cit. 2021-04-22]. Dostupné z: https://www.password-hashing.net/.
43. ADAMS, Carlisle. Dictionary Attack. In: *Encyclopedia of Cryptography and Security*. Ed. TILBORG, Henk C. A. van; JAJODIA, Sushil. Boston, MA: Springer US, 2011, s. 332–332. ISBN 978-1-4419-5906-5. Dostupné z DOI: 10.1007/978-1-4419-5906-5_74.
44. 1&1 IONOS INC. *Rainbow tables: Simply explained* [online]. 2021 [cit. 2021-06-12]. Dostupné z: https://www.ionos.com/digitalguide/server/security/rainbow-tables/.
45. AUTH0 INC. *Adding Salt to Hashing: A Better Way to Store Passwords* [online]. 2021 [cit. 2021-06-12]. Dostupné z: https://auth0.com/blog/adding-salt-to-hashing-a-better-way-to-store-passwords/.
46. CHEATSHEETS SERIES TEAM. *Session Management Cheat Sheet* [online]. 2021 [cit. 2021-12-13]. Dostupné z: https://cheatsheetseries.owasp.org/cheatsheets/Session_Management_Cheat_Sheet.html.
47. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *Using HTTP cookies* [online]. 2021 [cit. 2021-04-27]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/HTTP/Cookies.
48. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *Set-Cookie* [online]. 2021 [cit. 2021-04-27]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Set-Cookie.
49. MOZILLA AND INDIVIDUAL CONTRIBUTORS. *Types of attacks* [online]. 2021 [cit. 2021-04-27]. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/Security/Types_of_attacks.
50. NETSPARKER LTD. *Using the Same-Site Cookie Attribute to Prevent CSRF Attacks* [online]. 2021 [cit. 2021-06-12]. Dostupné z: https://www.netsparker.com/blog/web-security/same-site-cookie-attribute-prevent-cross-site-request-forgery/.

51. MONGODB, INC. *Types of Databases* [online]. 2021 [cit. 2021-12-27]. Dostupné z: https://www.mongodb.com/databases/types.
52. MONGODB, INC. *What is a Document Database?* [Online]. 2021 [cit. 2021-12-27]. Dostupné z: https://www.mongodb.com/it-it/document-databases.
53. MONGODB, INC. *Key-Value Databases* [online]. 2021 [cit. 2021-12-27]. Dostupné z: https://www.mongodb.com/databases/key-value-database.
54. ARANGODB, INC. *ArangoDB as Graph Database* [online]. 2021 [cit. 2021-12-27]. Dostupné z: https://www.arangodb.com/graph-database/.
55. WEINBERGER, Claudius. *Index Free Adjacency or Hybrid Indexes for Graph Databases* [online]. 2016 [cit. 2021-12-27]. Dostupné z: https://www.arangodb.com/2016/04/index-free-adjacency-hybrid-indexes-graph-databases/.
56. SOLID IT GMBH. *DB-Engines Ranking of Relational DBMS* [online]. 2021 [cit. 2021-12-27]. Dostupné z: https://db-engines.com/en/ranking/relational+dbms.
57. YIGAL, Asaf. *Sqlite vs. MySQL vs. PostgreSQL: A Comparison of Relational Databases* [online]. 2018 [cit. 2021-12-27]. Dostupné z: https://logz.io/blog/relational-database-comparison/.
58. THE POSTGRESQL GLOBAL DEVELOPMENT GROUP. *How Extensibility Works* [online]. 2021 [cit. 2021-12-27]. Dostupné z: https://www.postgresql.org/docs/9.0/extend-how.html.
59. ORACLE CORPORATION AND/OR ITS AFFILIATES. *MySQL 8.0 Reference Manual - Full-Text Search Functions* [online]. 2021 [cit. 2021-12-27]. Dostupné z: https://dev.mysql.com/doc/refman/8.0/en/fulltext-search.html.
60. BUILTWITH PTY LTD. *Web Server Usage Distribution in the Top 1 Million Sites* [online]. 2021 [cit. 2021-12-27]. Dostupné z: https://trends.builtwith.com/web-server.
61. ELLINGWOOD, Justin. *Apache vs Nginx: Practical Considerations* [online]. 2015 [cit. 2021-12-27]. Dostupné z: https://www.digitalocean.com/community/tutorials/apache-vs-nginx-practical-considerations.
62. GLEASON, Ryan. *Node.js vs. Spring Boot — Which Should You Choose?* [Online]. 2020 [cit. 2021-12-27]. Dostupné z: https://betterprogramming.pub/node-js-vs-spring-boot-which-should-you-choose-2366c2f76587.
63. APPLIED INFORMATICS SOFTWARE ENGINEERING GMBH. *POCO C++ LIBRARIES INFORMATION - FEATURES* [online]. 2021 [cit. 2021-12-27]. Dostupné z: https://pocoproject.org/about.html#features.
64. NGINX, INC. *NGINX* [online]. 2021 [cit. 2021-06-12]. Dostupné z: http://nginx.org/en/.

Obsah přiloženého média

- `readme.txt` ..... stručný popis obsahu média
- `thesis.pdf` ..... dokument ve formátu PDF
- `src`
  - `api` ..... zdrojové kódy implementace API (JS)
  - `db` ..... zdrojové kódy implementace databáze (SQL)
  - `data` ..... zdrojové kódy pro vygenerování insertscriptu (CSV, CPP, JS)
  - `tex` ..... zdrojová forma dokumentu (LaTeX)