

**Univerzita Palackého v Olomouci**  
**Přírodovědecká fakulta**  
**Katedra geoinformatiky**

**HISTORICKÉ TISKY VE VĚDECKÉ KNIHOVNĚ**  
**OLOMOUC**

**Magisterská práce**

**Vojtěch HAŠKA**

**Vedoucí práce doc. Mgr. Jiří DVORSKÝ, Ph.D.**

**Olomouc 2017**  
**Geoinformatika**

## **KLÍČOVÁ SLOVA**

Historické tisky; Vědecká knihovna Olomouc; dynamická webová mapa;

## **KEYWORDS**

Historical prints; Research library Olomouc; dynamic web map;

Počet stran práce: 35

Počet příloh: 2 (z toho 1 volná a 1 elektronická)

## **ANOTACE**

Tato diplomová práce se zabývá tvorbou dynamické webové mapy, která zobrazuje knihy z databáze Historických tisků Vědecké knihovny v Olomouci podle místa a času vydání.

## **ANOTATION**

This diploma thesis deals with creation of dynamic web map, that visualize books from repositories of Historical prints in Research library Olomouc according place and date of issue.

**Prohlašuji, že**

- diplomovou práci včetně příloh, jsem vypracoval samostatně a uvedl jsem všechny použité podklady a literaturu.

- při tvorbě aplikace jsem použil řešení pro zobrazování vektorových dat pomocí Leaflet a JQuery vytvořený na University of Wisconsin-Madison [5], rovněž jsem použil knihovnu pro sklizení OAI repozitářů od uživatele GitHub „kraenhasen“ [14] a XML parser od uživatele portálu StackOverflow „NexusRex“ [13]

- jsem si vědom, že na moji diplomovou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevýdělečně, ke své vnitřní potřebě, diplomovou práci užívat (§ 35 odst. 3),

- souhlasím, aby jeden výtisk diplomové práce byl uložen v Knihovně UP k prezenčnímu nahlédnutí,

- souhlasím, že údaje o mé diplomové práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé diplomové práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé diplomové práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne

Děkuji vedoucímu práce doc. Mgr. Jířimu Dvorskému, Ph.D. za podněty a připomínky při vypracování práce.

Za poskytnuté rady a materiály také děkuji zaměstnancům Vědecké knihovny v Olomouci.

## ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Vojtěch HAŠKA**  
Osobní číslo: **R120613**  
Studijní program: **N1301 Geografie**  
Studijní obor: **Geoinformatika**  
Název tématu: **Historické tisky ve Vědecké knihovně v Olomouci**  
Zadávající katedra: **Katedra geoinformatiky**

### Z á s a d y p r o v y p r a c o v á n í :

Cílem práce je zmapovat původ historických tisků ve Vědecké knihovně v Olomouci. V rámci diplomové práce vznikne webová aplikace, která bude na vhodně vybraném mapovém podkladu zobrazovat místa vzniku jednotlivých historických tisků.

Zobrazovaná data bude možno třídit podle data vzniku, časového období vzniku, země vzniku a dalších kritérií. Výsledná aplikace by měla sloužit k atraktivní prezentaci historických fondů VK v Olomouci.

Student vyplní údaje o všech datových sadách, které vytvořil nebo získal v rámci práce, do Metainformačního systému katedry geoinformatiky a současně zálohu údajů ve formě validovaného XML souboru. Celá práce (text, přílohy, výstupy, zdrojová a vytvořená data, XML soubor) se odevzdá v digitální podobě na CD (DVD) a text práce s vybranými přílohami bude odevzdán ve dvou svázaných výtiscích na sekretariát katedry. O diplomové práci student vytvoří webovou stránku v souladu s pravidly dostupnými na stránkách katedry. Práce bude zpracována podle zásad dle Voženílek (2002) a závazné šablony pro diplomové práce na KGI. Na závěr diplomové práce připojí student jednostránkové resumé v anglickém jazyce.

Rozsah grafických prací: **dle potřeby**  
Rozsah pracovní zprávy: **max. 50 stran**  
Forma zpracování diplomové práce: **tištěná**  
Seznam odborné literatury:

**Katalog historických forndů VK v Olomouci**  
**Mapové podklady**  
**Literatura ke zvolené webové technologii**

Vedoucí diplomové práce: **Doc. Mgr. Jiří Dvorský, Ph.D.**  
Katedra geoinformatiky


Datum zadání diplomové práce: **12. prosince 2012**

Termín odevzdání diplomové práce: **23. dubna 2014**

UNIVERZITA PALACKÉHO V OLOMOUCI  
PŘÍRODOVĚDECKÁ FAKULTA  
KATEDRA GEOINFORMATIKY  
17. listopadu 50, 771 46 Olomouc  
-1-

L.S.

Prof. RNDr. Juraj Ševčík, Ph.D.  
děkan

  
Prof. RNDr. Vít Voženílek, CSc.  
vedoucí katedry

V Olomouci dne 12. prosince 2012

# OBSAH

<b>SEZNAM POUŽITÝCH ZKRATEK .....</b>	<b>8</b>
<b>ÚVOD .....</b>	<b>9</b>
<b>1 CÍLE PRÁCE.....</b>	<b>10</b>
<b>2 METODY A POSTUPY ZPRACOVÁNÍ.....</b>	<b>11</b>
<b>SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY .....</b>	<b>13</b>
2.1 Webové mapy jako okna do databáze.....	13
2.2 Vznik a vývoj Historických fondů Vědecké knihovny Olomouc.....	14
2.3 Data v online repozitáři knihovny .....	15
2.4 Stažení dat ze serveru .....	15
2.4.1 OAI-PMH .....	15
2.4.2 Použití OAI-PMH .....	15
2.4.3 OAI Parser v PHP .....	15
2.4.4 Konverze z XML do databáze .....	16
2.5 Úprava a odstranění chyb v datech .....	16
2.5.1 Stav dat .....	16
2.5.2 Formát dat.....	16
2.5.3 Výběr hodnot určených k vizualizaci.....	16
2.5.4 Chyby v určení polohy, geokódování .....	17
2.5.5 Chyby v určení roku vydání.....	19
2.6 Vytvoření a naplnění databáze .....	20
2.7 Předávání dat v internetové aplikaci .....	20
2.7.1 SQL dotaz pro výběr měst.....	20
2.7.2 Výpis jednotlivých knih pro zadané období.....	23
2.8 Stažení ukázkových dat .....	24
<b>3 VIZUALIZACE .....</b>	<b>25</b>
3.1 Webové mapy.....	25
3.2 Historie webových map .....	25
3.3 Volba mapového podkladu .....	25
3.4 Volba vizualizační (kartografické) metody .....	26
<b>4 FUNKCIONALITA APLIKACE.....</b>	<b>26</b>
4.1 Rozložení ovládacích prvků .....	26
<b>5 VÝSLEDKY PRÁCE .....</b>	<b>29</b>
<b>6 DISKUZE .....</b>	<b>30</b>
<b>7 ZÁVĚR .....</b>	<b>31</b>
<b>POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE.....</b>	<b>34</b>
<b>PŘÍLOHY</b>	



## SEZNAM POUŽITÝCH ZKRATEK

<b>Zkratka</b>	<b>Význam</b>
CSV	Comma separated value
GeoJSON	Geographic JavaScript object notation
Marc21	Machine-Readable Cataloging
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
SQL	Structured Query Language
VKOL	Vědecká knihovna Olomouc
XML	Extensible markup language
UNGEGN	United Nations Group of Experts on Geographical names
AJAX	Asynchronous JavaScript and XML

## ÚVOD

V posledních desetiletích došlo k revolučním změnám ve správě a přístupu k informacím. Informace se začaly více ukládat a zpracovávat pomocí informačních technologií. Počítače se dostaly nejen do rukou profesionálů, ale i obyčejných lidí. V dnešní době produkujeme data, aniž bychom si to uvědomovali, skrze přístroje, které v dřívější době žádná digitální data neprodukovaly – např. automobily, fotoaparáty, telefony atd.

S rozvojem a rozšířením nových technologií souvisí rozvoj informační společnosti. Dnešní člověk je zvyklý každodenně informace přijímat a pracovat s nimi pomocí digitálních zařízení a internetu. Slovníky jsou vytlačovány internetovými vyhledávacími, knihy je možné číst v elektronické podobě a vyhledávat v nich, internet nabízí spojení s nepřeborným množstvím zdrojů, papírové mapy jsou dnes často nahrazovány těmi digitálními. Kdokoli se může podílet na tvorbě obsahu internetu. Množství dat neustále a rychle roste. Proto je kladen důraz na zlepšování struktury a informační hodnoty dat. To napomáhá snadnějšímu vyhledávání a orientaci v nich.

Strukturovaná data však sama o sobě nedokáží předat informaci, kterou chce člověk najít. Je třeba poskytnout uživateli nástroje, pomocí kterých se k nim dostane co nejjednodušší cestou a které mu poskytnou rozličné způsoby, jak data prohlížet. Vhodná vizualizace dat může výrazně urychlit zpracování informací příjemcem nebo zatraktivnit samotný zdroj dat pro uživatele. Mnoho dat v dnešní době má svůj prostorový aspekt, obsahují prostorové informace. Jedním z účelů geoinformačních technologií je data vizualizovat.

Někdo řekl, že obrázek vydá za tisíc slov. U mapy to platí dvojnásob. Mapová aplikace však dokáže ještě víc, protože se stává oknem do databáze. Jedním z datových zdrojů, který obsahuje prostorové informace a dosud se ho nepokusil nikdo zobrazit pomocí mapy je databáze metadat o knihách Vědecké knihovny v Olomouci, konkrétně databáze Historických tisků. Kromě prostorového aspektu mají tato data i aspekt časový. Jsou vhodné pro vytvoření dynamické mapové aplikace.

Protože jsou data volně přístupná na internetu, je záměrem autora, aby i aplikace byla dostupná pomocí internetu široké veřejnosti. Tato práce se zabývá tvorbou dynamické webové mapy, která zobrazuje knihy z databáze Historických tisků Vědecké knihovny v Olomouci podle místa a času vydání.

# 1 CÍLE PRÁCE

Cílem magisterské práce je vytvořit dynamickou webovou mapovou aplikaci zobrazující vhodným způsobem metadatové záznamy o historických tiscích uložených ve Vědecké knihovně v Olomouci s důrazem na jejich časový aspekt.

Mezi dílčí cíle patří: **získat data, upravit je, vybrat vhodnou metodu pro vizualizaci, naprogramovat aplikaci.**

Prvním cílem je získat data. Data jsou poskytována volně na serveru vkol.cz pomocí speciálního protokolu. Protože data obsahují chyby a jsou nekonzistentní, je třeba promyslet metody, jak se s těmito chybami vypořádat, aby bylo možné data v mapě správně zobrazit.

Dále je nutné vybrat vhodnou metodu vizualizace. Existují různé způsoby, jak z dat udělat mapu. Je třeba rozhodnout, která data bude možné vyhledávat pomocí vyhledávače.

Dalším cílem je vytvořit samotnou webovou aplikaci pomocí současných technologií pro vytváření webových aplikací a map. Výsledná práce umožní nový pohled na databázi Historických tisků pro její uživatele, ale také pro studenty historie a geografie. Vizualizace počtu knih v daném roce pomůže studentům získat informace o významu měst podle toho v jakých obdobích a v jakém množství se tu vydávaly knihy. Práce rovněž zvýší atraktivitu Vědecké knihovny v Olomouci a poskytne uživatelům nový pohled na knihy, které spravuje.

## **2 METODY A POSTUPY ZPRACOVÁNÍ**

Nejdříve bylo nutné zajistit způsob získání zdrojových dat pro tvorbu aplikace. Následně byla samotná data podrobně zkoumána, aby bylo možné rozhodnout, která část z nich bude dále použitelná. Byla promyšlena struktura aplikace, zvoleny vhodné technologie a nastudovány způsoby jejich užití.

### **Použité metody**

Při tvorbě práce bylo využíváno internetových i knižních zdrojů. Pro jednotlivé dílčí cíle byly vyhledávány fungující postupy a software. Následně byly postupy či software vhodně implementovány. Byly zkoumány aktuální způsoby a přístupy při tvorbě aplikací. Volba struktury aplikace, použití technologií, zpracování dat a použití vizualizačních prostředků byly konzultovány s odborníky.

### **Použitá data**

Použitá data mají svůj původ v databázi VKOL. Byla upravena a doplněna tak, aby byla lépe využitelná pro daný účel, ale zároveň aby nedošlo k jejich chybné úpravě a degradaci.

### **Použité programy**

Z webu byla data stažena pomocí OAI-PMH Harvester manageru 3.1.1. Překonvertována z formátu XML do formátu CSV pomocí nástroje XML2CSV. Data byla zpracována pomocí MS Excel 2010, studentské licence poskytnuté Univerzitou Palackého, dále byly použity skripty v jazyce Python 3.4 a Python 2.7.9. Data byla naimportována pomocí nástroje pro správu databází PHPMyAdmin 3.3.8 do databáze MySQL 5.1.

Později však byl tento postup přehodnocen a nahrazen více přímočarým řešením, které nevyžaduje tolik externích aplikací. Data jsou stahována pomocí OAI-PMH protokolu přímo ze serveru, přičemž skript je naprogramován v PHP. Data jsou uložena přímo v databázi MySQL.

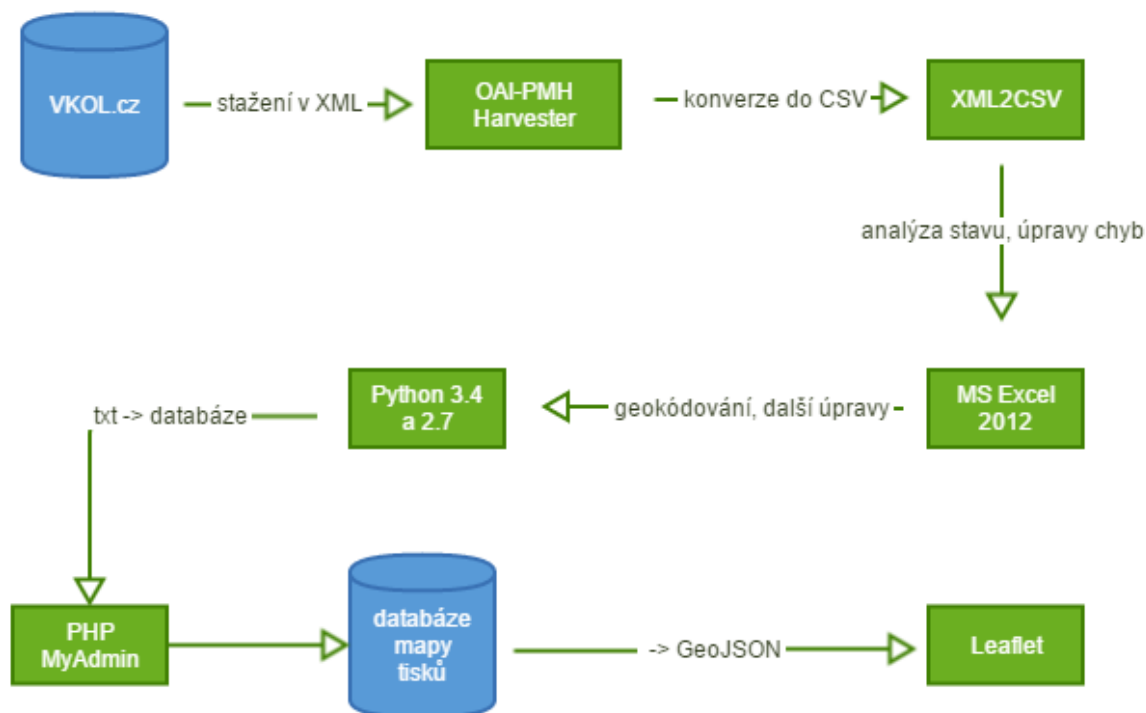
Pro vizualizaci byl použit framework Leaflet 0.7.2.

### **Použité technologie a programovací jazyky**

Při zpracování dat bylo využíváno skriptů v jazyce Python 2.7 a 3.5. Byly použity vzorce a makra v programu MS Excel. Pro naprogramování samotné aplikace na straně serveru byl využit jazyk PHP, pro komunikaci se serverem a zobrazování mapy pak AJAX, JQuery, JavaScript, CSS a HTML.

### **Postup zpracování**

Postup zpracování dat popisuje diagram na obrázku níže. Nejprve byla data stažena z databáze vkol.cz pomocí OAI-PMH Harvesteru, poté byla konvertována do CSV prostřednictvím programu XML2CSV, aby mohla být následně analyzována v tabulkovém procesoru MS Excel 2012. V programu Excel byly provedeny analýzy četností výskytu záznamů podle jednotlivých tagů, byly zde rovněž vytvořeny šablony pro geokódování a doplnění některých chybějících hodnot.



Obr. 1 Schéma popisující cestu zpracování dat od z repozitáře VKOL.cz až po vizualizaci pomocí frameworku Leaflet.

Následně byla data zpracovávána prostřednictvím Python (verze 3.5) skriptů. Python byl vybrán kvůli velkému objemu dat a potřebám práce s daty, pro které již vzorce v Excelu nebyly dostatečné. Pomocí Python skriptů bylo dokončeno geokódování a byly vyřazeny nadbytečné a chybné hodnoty. Výstupem z Python skriptů a Excelu jsou textové soubory obsahující hodnoty vždy jednoho z tagů, které jsou upraveny pro použití v aplikaci, mají jednotné názvy míst a roků vydání bez nadbytečných znaků a s menší pravděpodobností dvojnáčnosti zejména u míst vydání.

Takto připravená data jsou pomocí skriptů `add_book.php` a `add_city.php` vloženy do databáze MySQL. Skripty `book_list.php` a `city-data.php` byly vytvořeny, aby vytáhly potřebná data z databáze a upravily je do formátu, který bude zobrazitelný pomocí frameworku Leaflet a tabulku v pravém menu podle zadaných parametrů uživatele.

## **SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY**

V současné době na stránkách Vědecké knihovny neexistuje podobná webová mapa. Ani jiné knihovny v České Republice ani v zahraničí nespustily žádný podobný projekt. Záznamy o knihách jsou stále dostupné pouze v databázích nebo tabulkách.

### **2.1 Webové mapy jako okna do databáze**

Webová kartografie je označována jako trend již značnou dobu. V současnosti přibývají nové trendy v rámci této disciplíny tvorby map. V kontextu prostorového zpracování dat, je proces kartografické vizualizace považován za překlad nebo konverzi prostorových dat z databáze do jakoby mapových produktů. Tento proces je veden slovy „Jak něco někomu říci a jak to udělat efektivně?“ Rozvinutí tohoto principu dalo slovu vizualizace širší význam. Podle slovníku to znamená „učinit viditelným“ a není pochyb o tom, že toto bylo úkolem kartografů vždy. Nicméně, pokrok v dalších disciplínách dovedl toto slovo do mnohem konkrétnějších způsobů, které mohou moderní počítačové technologie popsat jako proces „učinit viditelným“ v reálném čase. [1]

V minulosti hrála kartografie důležitou roli při prozkoumávání světa. Mapy byly používány pro načrtnutí neznámých teritorií. Nová fáze mapování neznámého začala právě nedávno. A ta nemá nic společného s kartografickým nebo geografickým průzkumem. Toto mapování je mapováním kyberprostoru. [1]

Na tyto aktuální výzvy se snaží poukázat a reagovat tato diplomová práce. Metadatové záznamy starých tisků ve Vědecké knihovně Olomouc jsou sice dostupné v elektornické podobě a kdokoli v nich může vyhledávat nebo, pokud je dokáže stáhnout, může je zobrazit ve formě tabulek. Z pohledu nových trendů v geoinformaticce je to však dosud nezmapované místo v kyberprostoru, na které může být tímto způsobem poukázáno. Prostřednictvím vizualizace dat, které byly dříve dostupné pouze ve formě tabulek tak vzniká nový pohled na databázi. Tento pohled může oslovit a inspirovat odborníky i laiky, kteří se zabývají historií, geografii nebo starými knihami. Tito lidé mohou být odborníky na svůj obor a mohou mít spoustu zkušeností. Často však nemají takovou informační a geoinformační gramotnost aby si dokázali uvědomit stále přibývající nové možnosti, které jsou v současnosti dostupné díky prudkému rozvoji informačních technologií a digitalizaci.

## 2.2 Vznik a vývoj Historických fondů Vědecké knihovny Olomouc

Počátky i pozdější vývoj tzv. historických knižních fondů dnešní Vědecké knihovny v Olomouci jsou úzce spjaty s dějinami olomoucké univerzity. V roce **1566** je v Olomouci zřízena jezuitská akademie. Ještě téhož roku započala práce na vybudování knihovny. Protože se roku 1573 stala jezuitská akademie plnohodnotným vysokým učením, musela této hodnosti odpovídat i kvalita knihovny – vzrostly nároky na její neustálé rozšiřování a správu. [7]

V roce 1599 došlo k nákupu většího počtu knih a knihovna zabírala tři místnosti. K velké ráně však došlo během třicetileté války. Osmiletý pobyt Švédů (1642-1650) v Olomouci a jejich zájem zejména o umělecké předměty, tedy i knižní sbíry způsobil nevratné a těžké škody. Údajně na sto povozech byly odevezeny spousty knih, kterými se dnes může chlubit Stockholmská knihovna.

Prvním dokumentem, který nás může informovat o seznamu knih je svazkový katalog z roku 1704 řazený podle oborů a autorů. V tomto roce se v knihovně nacházelo 4000-5000 svazků plus dalších 2000-2500 se nacházelo různě po ubikacích vyučujících apod. [7]

Po zrušení jezuitského řádu (1773) zůstalo v knihovně jen asi 8100 knih. Knihovna se poté roku 1775 stává veřejnou univerzitní knihovnou a byla významně obohacena o vybrané knižní fondy pěti moravských zrušených jezuitských kolejí. O pouhé tři roky později tak spravuje téměř 40 000 tiskopisů. Dalšímu významnému obohacení se mohla knihovna těšit po rozhodnutí císaře Josefa II. z roku 1781 a 1782, kdy zrušil všechny kláštery, které podle jeho soudu nebyly užitečné. Olomoucká knihovna se pak měla postarat o všechny knihy ze zrušených církevních institucí. Ačkoliv měla přednostní právo výběru knih knihovna ve Vídni, toto svoje právo příliš nevyužívala. Bylo nutné vybrat z 88000 děl v cca 400000 svazcích. Jan Alois Hanke však mnoho tisků zbytečně označil jako odpad a nakonec byla knihovna obohacena o jen asi 6000 děl.

V 19. století rovněž přibývalo knih díky pozůstalostem, darům a konfiskacím jiných fondů. K navýšení historických fondů došlo rovněž v 50. letech 20. století, kdy se též rušilo mnoho církevních institucí. Poslední větší nárůst byl spojen naopak s restitucemi v roce 1989. Posledním větším obohacením historických fondů bylo 5295 svazků z litoměřického kláštera dominikánů. [7]

Dnes tedy tvoří historické fondy **1 455 rukopisů** (z nich 401 středověkých), téměř **2 000 prvotisků** a **přes 65 000 starých tisků**, čímž se historické fondy Vědecké knihovny v Olomouci řadí mezi největší a nejvzácnější v České republice. [7]

## 2.3 Data v online repozitáři knihovny

Data jsou základním prvkem mapové aplikace. Data, která jsou ke stažení přímo ze serveru VKOL však nebylo možné přímo použít, a to ze tří hlavních důvodů:

1. nejsou ve vhodném formátu,
2. bylo třeba z nich vybrat pouze informace, které budou vhodné k použití,
3. bylo třeba upravit chybné položky, případně vyřadit nepoužitelné položky z databáze.

Aby bylo možné zajistit plynulé fungování mapové aplikace, je třeba stáhnout jednou za čas všechna a nemuset je už stahovat při každém požadavku uživatele, který navštíví stránku s webovým portálem.

## 2.4 Stažení dat ze serveru

Vědecká knihovna Olomouc poskytuje na svém webu pokročilé vyhledávací služby pro všechny uživatele. Pro hromadné stažení dat nabízí možnost využít protokolu OAI-PMH (Open Archive Initiative - Protocol for Metadata Harvesting).

### 2.4.1 OAI-PMH

**OAI** je iniciativa, která vyvíjí a prosazuje výměnné standardy s cílem umožnit efektivní zpřístupnění obsahu velkých online repozitářů. [11]

#### **OAI-PMH**

- je „protokol pro sklizení metadat, definuje mechanismus pro sklizení metadatových záznamů,
- poskytuje jednoduchý technický prostředek poskytovatelům dat pro účely zpřístupnění svých metadat službám založeným na obecně rozšířených standardech HTTP (Hypertext Transport Protocol) a XML (Extensible Markup Language).“ [12]

### 2.4.2 Použití OAI-PMH

OAI-PMH Adresa OAI serveru pro VKOL je <http://aleph.vkol.cz/OAI> [10], pro práci s protokolem se používají speciální slovesa, jako například: GetRecord, Identify, ListIdentifiers, ListMetadataFormats, ListRecords, ListSets, ap.

### 2.4.3 OAI Parser v PHP

Pro stahování dat z repozitářů existují různá softwarová řešení. Pro jednodušší následné aktualizace byl pro účely této práce vybrán přístup, pomocí kterého bude možné aktualizovat data přímo ve webové aplikaci. Aplikace, která data stáhne z repozitáře [vkol.cz](http://vkol.cz) je implementovaná na serveru a je naprogramována v jazyce PHP.

PHP parser uložený v souboru s názvem `marc.php` poskytuje základní nástroje pro práci s OAI-PMH. Pomocí této knihovny byl vytvořen nástroj pro zálohování dat určených pro vizualizaci. Stažení celého repozitáře VKOLOAI může trvat až 3 dny, pokud dojde k přerušení, je nutné spustit skript `marc.php` znovu a vložit na konec URL parametry uložené v souboru `text.txt`. Tyto parametry zajistí pokračování plnění databáze v místě, kde došlo k chybě a není nutné tak začínat stahovat celý repozitář znovu.



## 2.4.4 Konverze z XML do databáze

Data jsou zpřístupňována ve formátu XML. PHP skript stahuje obsah repozitáře po částech a vkládá data do připravené databáze v MySQL.

Pomocí cyklů for jsou probírány jednotlivé nody XML souboru a u těchto je zjišťováno zda obsahují informace podstatné pro vizualizaci.

## 2.5 Úprava a odstranění chyb v datech

Metadata, která popisují staré tisky uložené ve Vědecké knihovně byla původně vytvářena mnohem dříve, než existovaly počítače. Knihovna spravuje knihy staré více než 500 let a vydané v různých jazycích a místech.

Vzhledem k těmto skutečnostem se dá jen těžko předpokládat, že data budou připravena na okamžité strojové zpracování a nasazení do jednotného systému. Následující kapitola popisuje, jakým způsobem byla data zpracována a připravena pro vizualizaci.

### 2.5.1 Stav dat

Data vykazují značnou dávku chybovosti a v některých případech i nejistoty. Vypořádat se s těmito chybami a pokusit se o jejich odstranění bylo jedním z dílčích úkolů při realizaci výsledné práce.

Ačkoliv jsou data digitální a dalo by se předpokládat, že by měla být uložena podle klíčů určených standardy, při práci s nimi je možné narazit na spoustu netypických nebo mírně pozměněných, nebo zcela mylných záznamů.

Chyby můžeme rozdělit na hrubé, soustavné (systematické) a náhodné. Vzhledem k tomu, že odstranit hrubé chyby je z povahy studovaného předmětu nad rámec této práce, zaměřil se autor na vypořádání se s chybami zejména systematickými, případně náhodnými.

### 2.5.2 Formát dat

Knihovna poskytuje data ve dvou standardních knihovnických formátech – Marc21 a DublinCore. Formát Marc21 poskytuje strukturovaný pohled na metadata o knihách. Informace jsou uloženy v jednotlivých záznamech označených jako tagy a pojmenovaných pomocí čísel, nebo kombinací čísel a písmene, každá tato kombinace má svůj konkrétní význam. Kompletní výčet a význam všech tagů [6] byl zdrojem důležitým pro správné rozklíčování jednotlivých informací v datech a k určení informací vhodných pro další zpracování a využití při tvorbě aplikace. K datům je možno přistupovat ve formátu XML.

### 2.5.3 Výběr hodnot určených k vizualizaci

Pro vizualizaci metadatových záznamů bylo nejprve nutné vymezit jaká data budou nezbytná pro samotnou vizualizaci jednotlivých záznamů či zobrazení základní popisných informací a práci s dynamickým aspektem.

Protože výsledkem této práce je webová dynamická mapa, kromě aspektu prostorového je nezbytné, aby hledané záznamy měly korektní údaje rovněž v rámci aspektu časového.

Tab. 1 Výpis tagů vhodných pro vizualizaci v mapě

atribut	počet	výskyt v procentech	popis
1	94935	100	ID
8	94935	100	rok vydání, místo vydání
245a	94917	99,98104	název
260c	94281	99,31111	rok vydání
260a	93751	98,75283	místo vydání (zpravidla endonyma)
100a	73100	77,00005	název autora
984a	61961	65,26676	místo vydání (česká exonyma)
984b	61953	65,25833	země vydání (česká exonyma)

## 2.5.4 Chyby v určení polohy, geokódování

Endonymum je podle definice UNGEGN (United Nations Group of Experts on Geographical names) [8] oficiální standardizovaný název užívaný na území, na kterém se geografický objekt nachází, bez ohledu na písmo, v němž je vyjádřen v písemné podobě.

Domácí endonyma jsou názvy objektů ležících na území vydavatele a uživatele mapy uváděná v úředním jazyce daného státu;

Tzv. cizí endonyma, jsou jména mimo území, kde se mapa vydává a užívá.

Exonymum, tzv. vžitě jméno je geografický název používaný v určitém jazyce pro objekt na území mimo oficiální působnost tohoto jazyka, který má odlišnou podobu od názvu použitého v úředním jazyce (úředních jazycích) území, na němž se objekt nachází.

Každá členská země OSN je povinna vydat seznam exonym nacházejících se mimo její státní území. V České Republice má tento úkol na starost Návoslovná komise Českého úřadu zeměměřického a katastrálního. Rovněž administrativně a technicky zabezpečuje databázi geografických jmen Geonames.

Tabulka níže ukazuje rozličné údaje o poloze, které je možné nalézt v datech. Informace o poloze byly extrahovány zejména z tagů 8, 260a, 984a a 984b.

Protože tag 260c obsahuje informace o místě vydání téměř u každého záznamu bylo by ideální, kdyby šlo podle obsahu tohoto tagu jednoznačně určit místo vydání. Kvalita záznamů v rámci tohoto tagu je však velmi nízká.

K upřesnění pomůže tag 984a, kde je název místa vydání zpravidla zapsán korektně, v případě měst mimo ČR jsou tu zapsána zpravidla česká exonyma. Zde však je četnost záznamů podstatně nižší. Autor práce postupoval tak, že vytvořil klíčovací tabulku na základě záznamů, které měly vyplněny tagy 260c i 984a (případně i 984b). Na základě této tabulky byl vytvořen klíč, který umožňuje automatické přiřazení strojově čitelných názvů na základě podobnosti.

Např. pokud existuje název „[À Paris :“ který není evidentně strojově čitelný (není možné provést úspěšné geokódování), ale v tagu 984a můžeme najít „Paříž“, můžeme přiřadit „Paříž“ i do všech ostatních záznamů, v kterých můžeme sice najít „[À Paris :“ v tagu 260c, v tagu 984a je však jen chybějící hodnota. Tímto způsobem bylo možné o něco navýšit počet záznamů, které je možné umístit do mapy, protože exonymum „Paříž“ je možné úspěšně umístit do mapy pomocí dostupných metod geokódování.

Další možností jak přispět ke snazšímu geokódování bylo odstranit z hodnot v tagu 260c znaky, které se běžně v názvech měst nevyskytují – např „!“ , „=“ , „?“ , nebo „:“. Tato úprava rovněž přispěla k četnosti záznamů, které bylo možné do mapy umístit.

Tab. 2 Stav zdrojových dat popisujících místa vydání, ukázka z dat.

008	260c	984a	984b
030318s1788 gw f ger	Leipzig :	Lipsko	Německo
030318s1759 fr f fre	En Europe[=Paris,] :	Paříž	Francie
030318s1592 sz f lat	[Geneva] :	Ženeva	Švýcarsko
030318s1679 fr f fre	[À Paris :	Paříž	Francie
030318s1795 au f fre	A Vienne :	Vídeň	Rakousko
030318s1493 gw f lat	Lipsiae :	Lipsko	Německo
030318s1662 sp f spa	Madrid (Por Joseph Fernandez de Buendia) :	Madrid	Španělsko
030318s1679 fr f lat	Lugduni :	Lyon	Francie
030319s1572 gw f lat	Coloniae Agrippinae :	Kolín nad Rýnem	Německo
030319s15011800xx f ger	[S.l. :	---	---
030319s1570 xr f lat	Pragae :	Praha	Česká Republika
030320s1890 xr f 000 j cze	V Praze :	---	---
030320s1618 ne f lat	Dordraci :	Dordrecht	Nizozemsko
990630s1820 hu f ger	Ungarisch-Altenburg,	---	---
010611s1628 xx f ger	[S.l. :		
010701s1840 au	Vienne :	---	---
010605s1679 xr f lat	W Holomaucy :	Olomouc	Česko
010606s1796 gw f ger	[S.l. :		
010607s1737 xx f fre	A la Haye :	Haag	Nizozemsko
060220s1845 xr f cze	Znogmo :		
060220s1888 xr f cze	V Čáslavi :		
060220s1855 xr f ger	Sternberg :		
060220s1856 xr f ger	Brünn :		
060220s1882 xr f pol	Cieszyn :		
060221s1858 xr f ger	Teschen :		
060221s1875 xr f lat	Cremsirii :		
060221s1833 xr f pol	Berno :		
060221s1885 xr f cze	V Brně :		
060221q1833 xr f ger	Olmütz :		
060301s1898 xr e cze	Mor. Ostrava :		

## 2.5.5 Chyby v určení roku vydání

Chyby v zadání míst vydání byly značné. V případě doby vydání jich nebylo méně. Jak je možné vidět v záznamech níže, chyby určení doby jsou rovněž značné a rozmanité. Od absence jakéhokoliv zadání v tagu 260c přes přebytné znaky a slova, výrazy označující nejistotu (asi, ?, kolem) až po chyby vzniklé pravděpodobně chybným zadáním do samotného systému.

Tab. 3 Chyby časových údajů, ukázka z dat.

008	260c
030318s1662 sp f spa	
030318s1493 gw f lat	1493
030318s1759 fr f fre	[1759]
030319q15011800au f ger	[16.-18. st.]
030319s15011800xx f ger	16.-18. st.]
030319s1893 au f cze	1893:00:00
990630s1890 xr d 00 j cze	[kolem 1890]
990630q1890 xr f cze	[ca. 1890]
990630s1892 xr d 00 p cze	[asi 1892]
990630q18611862xr f cze	1861-1862
050825q18821886xr abf g 001 0 cze	[1882?]
990630s1885 xr f	[kolem r. 1885]
990630s1849 xr f ger	1849?]
990630s1874 gw	[mezi 1874-79]
990630s1890 gw	c1890
990630m18501860xr f zxx	[1850-1860]
990630s188 ru f	188
990630q18uu xr f 000 ger	[18--]
990630s1850 xr f cze	[kolem 1850]
990630s1887 gw f	[před 1887]
010320q18001850xr f cze	[mezi 1800-1850?]
020211m18351889gw f lat	1835-1889
020211s18261896gw f lat	1826-1896
990630s185 xr f	185
030208s1867 xr f	vl. n. 1867
010701s181 xr f	1981

## 2.6 Vytvoření a naplnění databáze

Databáze určená pro potřeby aplikace nemusí přesně kopírovat údaje které se nachází na serverech Vědecké knihovny. Jejím účelem je zejména vizualizovat záznamy

a poskytnout základní informace. Pokud bude uživatel chtít, může si podrobnější data vyhledat sám u jednotlivých záznamů. U každého záznamu může přejít přímo na webovou stránku Vědecké knihovny a získat tak kompletní detailní informace o zvolené knize.

Pro vytvoření databáze na serveru slouží připravený soubor SQL s názvem „struktura\_db.sql“ ve složce import\_export, který postačí importovat do databáze s názvem „print“. Tento SQL soubor vytvoří potřebnou hierarchii. Pro naplnění databáze daty knih a měst je třeba spustit skripty v souborech add\_book.php a add\_city.php.

## 2.7 Předávání dat v internetové aplikaci

Samotná aplikace používá databázi MySQL. Data z ní jsou extrahována pomocí jednotlivých SQL příkazů. Tyto příkazy jsou formulovány pomocí PHP skriptu v reálném čase. Pro spuštění samotných skriptů je využita technologie AJAX (asynchronní javascript a xml). Tato technologie umožňuje plynulejší práci s daty. Jednou z výhod použití technologie AJAX je, že uživatel často ani nezpozoruje, že se data načítají ze serveru, protože se kvůli načítání dat nemusí znovu obnovovat celá webová stránka, ale mění se pouze její část, čímž dochází rovněž ke značnému zrychlení.

Aby bylo možné data vizualizovat pomocí frameworku Leaflet, musí být zkonvertována do formátu JSON. PHP skript je nastaven tak, aby data v tomto formátu produkoval.

### 2.7.1 SQL dotaz pro výběr měst

Hlavním tematickým prvkem celé mapy jsou města, která jsou reprezentována bodovými znaky měnicími velikost podle počtu knih vydaných v daném městě za zvolené období.

Aby byla zaručena plynulost aplikace a hostingový server nebyl přetěžován, načítají se data do mezipaměti pouze pokud uživatel změní interval, který určuje za jak dlouhé období se mají knihy počítat.

SQL příkaz tedy vybírá všechna města z tabulky, u kterých byly určeny zeměpisné souřadnice a kde byla vydána v jakémkoli období alespoň jedna kniha. Tento příkaz vrací tabulku, v které je pro každé město určena jednoznačná poloha a každý interval ve zvoleném období zde má vypsán počet knih. Na základě těchto dat, která jsou převedena do formátu JSON skript zobrazí výslednou mapu.

Následující kód představuje SQL příkaz, který se uskuteční, když uživatel požaduje zobrazení mapy historických tisků s intervalem 10 let.

```

select period, id, lat, lng, MAX(count) AS count
from (      select * from      (select t4.i*1000 + t3.i*100 + t2.i*10 +
t1.i period
from      (select 0 i union select 1 union select 2 union select 3 union
select 4 union select 5 union select 6 union select 7 union select 8
union select 9) t1,
(select 0 i union select 1 union select 2 union select 3 union select 4
union select 5 union select 6 union select 7 union select 8 union select
9) t2,
(select 0 i union select 1 union select 2 union select 3 union select 4
union select 5 union select 6 union select 7 union select 8 union select
9) t3,
(select 0 i union select 1 union select 2 union select 3 union select 4
union select 5 union select 6 union select 7 union select 8 union select
9) t4) numbers

inner join city inner join (select 0 count) zeros
where period >= 1400 and period <= 1919 and period mod 10 = 0
AND city.lat IS NOT NULL AND city.lng IS NOT NULL
union SELECT floor(book.year / 10) * 10 as period,      city.id,
city.lat, city.lng, COUNT(*) as `count`      FROM book, city
WHERE book.id_city = city.id      GROUP BY period, id_city      HAVING
period IS NOT NULL AND city.lat IS NOT NULL AND city.lng IS NOT NULL
AND period >= 1400 AND period <= 1919      ) valid_rows      GROUP BY id,
period      ORDER BY id, period

```

Tento SQL příkaz je zavolán do databáze, která vrací konkrétní data pro všechna města a všechna možná období ve výběru. Pokud pro dané období a dané město nejsou nalezeny žádné záznamy, této příkaz přiřadí hodnotu nula. Pokud město nemá přiřazené hodnoty nutné k lokalizaci, nezobrazí je. Na obrázku níže je ukázkový výpis, který vrací databáze po zavolání výše uvedeného SQL příkazu, který je vygenerován s hodnotou intervalu 10 let.

1400	Augsburg	48.370544433593750	10.897789955139160	0
1410	Augsburg	48.370544433593750	10.897789955139160	0
1420	Augsburg	48.370544433593750	10.897789955139160	0
1430	Augsburg	48.370544433593750	10.897789955139160	0
1440	Augsburg	48.370544433593750	10.897789955139160	0
1450	Augsburg	48.370544433593750	10.897789955139160	0
1460	Augsburg	48.370544433593750	10.897789955139160	0
1470	Augsburg	48.370544433593750	10.897789955139160	27
1480	Augsburg	48.370544433593750	10.897789955139160	23
1490	Augsburg	48.370544433593750	10.897789955139160	36
1500	Augsburg	48.370544433593750	10.897789955139160	25
1510	Augsburg	48.370544433593750	10.897789955139160	42
1520	Augsburg	48.370544433593750	10.897789955139160	79
1530	Augsburg	48.370544433593750	10.897789955139160	34
1540	Augsburg	48.370544433593750	10.897789955139160	19
1550	Augsburg	48.370544433593750	10.897789955139160	9
1560	Augsburg	48.370544433593750	10.897789955139160	6
1570	Augsburg	48.370544433593750	10.897789955139160	7
1580	Augsburg	48.370544433593750	10.897789955139160	9
1590	Augsburg	48.370544433593750	10.897789955139160	15
1600	Augsburg	48.370544433593750	10.897789955139160	14
1610	Augsburg	48.370544433593750	10.897789955139160	20
1620	Augsburg	48.370544433593750	10.897789955139160	12
1630	Augsburg	48.370544433593750	10.897789955139160	11
1640	Augsburg	48.370544433593750	10.897789955139160	9
1650	Augsburg	48.370544433593750	10.897789955139160	10
1660	Augsburg	48.370544433593750	10.897789955139160	10
1670	Augsburg	48.370544433593750	10.897789955139160	17
1680	Augsburg	48.370544433593750	10.897789955139160	19
1690	Augsburg	48.370544433593750	10.897789955139160	64
1700	Augsburg	48.370544433593750	10.897789955139160	128
1710	Augsburg	48.370544433593750	10.897789955139160	187
1720	Augsburg	48.370544433593750	10.897789955139160	273
1730	Augsburg	48.370544433593750	10.897789955139160	205
1740	Augsburg	48.370544433593750	10.897789955139160	102
1750	Augsburg	48.370544433593750	10.897789955139160	172
1760	Augsburg	48.370544433593750	10.897789955139160	277
1770	Augsburg	48.370544433593750	10.897789955139160	265
1780	Augsburg	48.370544433593750	10.897789955139160	147
1790	Augsburg	48.370544433593750	10.897789955139160	93
1800	Augsburg	48.370544433593750	10.897789955139160	7
1810	Augsburg	48.370544433593750	10.897789955139160	0
1820	Augsburg	48.370544433593750	10.897789955139160	0
1830	Augsburg	48.370544433593750	10.897789955139160	0
1840	Augsburg	48.370544433593750	10.897789955139160	0
1850	Augsburg	48.370544433593750	10.897789955139160	0
1860	Augsburg	48.370544433593750	10.897789955139160	0
1870	Augsburg	48.370544433593750	10.897789955139160	0
1880	Augsburg	48.370544433593750	10.897789955139160	0
1890	Augsburg	48.370544433593750	10.897789955139160	0
1900	Augsburg	48.370544433593750	10.897789955139160	0

Obr. 2 Databáze vypisuje údaje potřebné k vizualizaci. Data jsou následně konvertována do formátu GeoJSON.

## 2.7.2 Výpis jednotlivých knih pro zadané období

Následující blok obsahuje zdrojový kód naprogramovaný v PHP uložený v souboru `book_list.php`, který je volán pomocí asynchronního příkazu. Komentáře u zdrojového kódu popisují funkcionality jednotlivých oddílů zdrojového kódu.

```
<?php
date_default_timezone_set('Europe/Prague');
header('Content-Type: application/json');
include "db.php";
//vypsání konkrétního seznamu knih na základě zadaných kritérií
(slova vyskytující se v metadatovém záznamu, rok vydání od, do,
konkrétní místo vydání)
if (isset($_GET['search'])) $search = $_GET['search']; else $search
= '';
if (isset($_GET['from'])) $from = $_GET['from']; else $from =
0;
if (isset($_GET['to'])) $to = $_GET['to']; else $to =
3000;
if (isset($_GET['limit'])) $limit = $_GET['limit']; else $limit =
0;
if (isset($_GET['place'])) $place = $_GET['place']; else $place =
'';

function stringParts ($text, $delimiter = ':')
{
    $array = preg_split("/\\\\".$delimiter."/",$text);
    return $array;
}
```

Následující část kódu popisuje konstrukci SQL dotazu pro dané město a zvolený interval, přidává uživatelem zvolené další klíčové parametry pro vyhledávání. Tato klíčková slova se prohledávají ve všech sloupcích databáze. Je tedy možné do fulltextového vyhledávače zadat jak rok, tak město, název titulu, nebo jméno autora. Uživatelem zadané heslo bude systém hledat ve všech sloupcích a všech záznamech.

```
function constructWhereQuery($userQuery, $from, $to, $place)
{
if ($userQuery=="IMPORT ALL") $q="IMPORT ALL";
else
    $parts = stringParts($userQuery, ' ');
    $q = "year >= $from ";
    $q .= "AND year <= $to ";
    if ($place!='') { $q .= "AND id_city = '$place.'" "; }

foreach ($parts as $part)
{
```



```

        $q .= " AND (id_original LIKE '%" . $part . "%' OR year LIKE
 '%" . $part . "%' ".
        "OR author LIKE '%" . $part . "%' OR title LIKE '%" . $part . "%' OR
id_city LIKE '%" . $part . "%') ";
    }

    return $q;;
}

function constructLimitQuery($userQuery, $userLimit)
{
    if ($userQuery == '') return 'LIMIT 0';
    if ($userLimit == 0) return;
    return 'LIMIT ' . $userLimit;
}

//vyhledávací dotaz
$sql = "SELECT * FROM book WHERE ". constructWhereQuery($search,
$from, $to, $place)." ". constructLimitQuery($search.$place, $limit);
$result = $conn->query($sql);

if (!$result) //selhání dotazu
{
    print 'QUERY: ' . $sql. "<br/>";
    print '<strong>ERROR: ' . $conn->error. "</strong><br/>";
}

//přečti výsledky
$book = array();
while ($row = $result->fetch_assoc())
{
    $book[] = $row;
}
print json_encode($book);

```

## 2.8 Stažení ukázkových dat

Pro uživatele, kteří by chtěli data kompletně stáhnout je možné přejít do složky `./import_export/index.html`. Zde je možné stáhnout všechna data ve formátu GeoJSON a to jak pro jednotlivá města tak pro každou knihu. Data potom může uživatel zpracovávat pro svou vlastní potřebu např. v desktopovém programu.

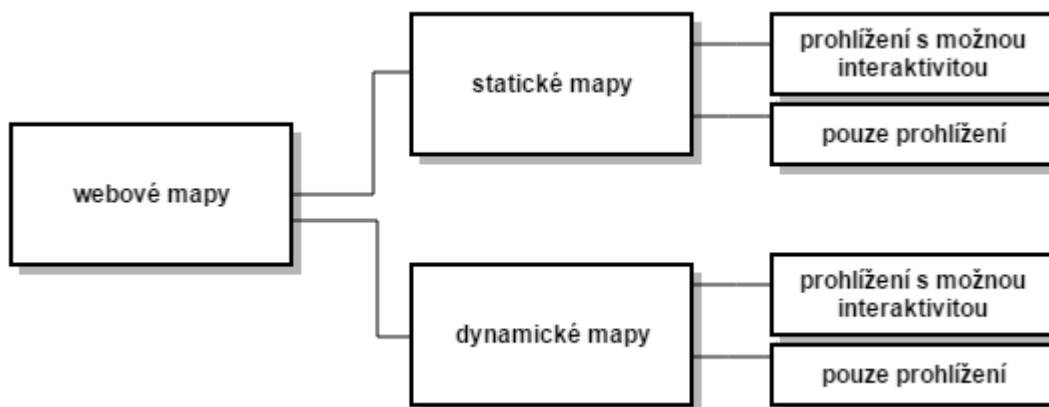
## 3 VIZUALIZACE

### 3.1 Webové mapy

Mapy webové, mají oproti offline digitálním a papírovým mnoho nesporných výhod. Webová mapa je přístupná odkudkoli v jakýkoli okamžik ve stejné kvalitě. Je možné ji přibližovat, měnit nastavení zobrazení, měnit informace, které jsou na ni zobrazené a spoustu dalšího, je možné ji též použít jako okno do databáze. Papírové mapy jsou oproti tomu v zásadě statické.

### 3.2 Historie webových map

S vynálezem webu a spuštění veřejné sítě internet jako takové v roce 1989 bylo spjato mnoho očekávání do budoucna. Netrvalo to dlouho a vedle prvních vět, článků, obrázků a jiných informačních entit přišly i první online webové mapy. Zpočátku mapy statické, vznikly jejich první prohlížečky, postupně přibýly mapy dynamické. Mapy dnes můžeme dělit podle Kraaka a Browna na statické a dynamické, dále v každé z těchto kategorií na mapy pouze k prohlížení a mapy interaktivní.[1]



Obr. 3 klasifikace webových map (podle Kraaka a Browna)

### 3.3 Volba mapového podkladu

Podkladová mapa byla vložena pomocí služby MapBox. Pro její zprovoznění a běh stačí bezplatná verze. [3][2] Podkladová mapa by měla vhodně doplňovat hlavní tematickou část mapy. Byla vybrána tak aby splňovala prvek atraktivity, zároveň však aby příliš nezatěžovala obsah nadměrnými informacemi, které nejsou pro daný účel podstatné.

### 3.4 Volba vizualizační (kartografické) metody

Data, s kterými tato práce pracuje, jsou diskrétní. Každý záznam reprezentuje jeden výtisk jedné konkrétní knihy, která má svoje unikátní místo vydání i rok vydání. Předpokládá se, že v jednom městě může být v určitém roce vydaných hned několik knih. Jednotlivá města budou představovat různé kvantify v různých letech.

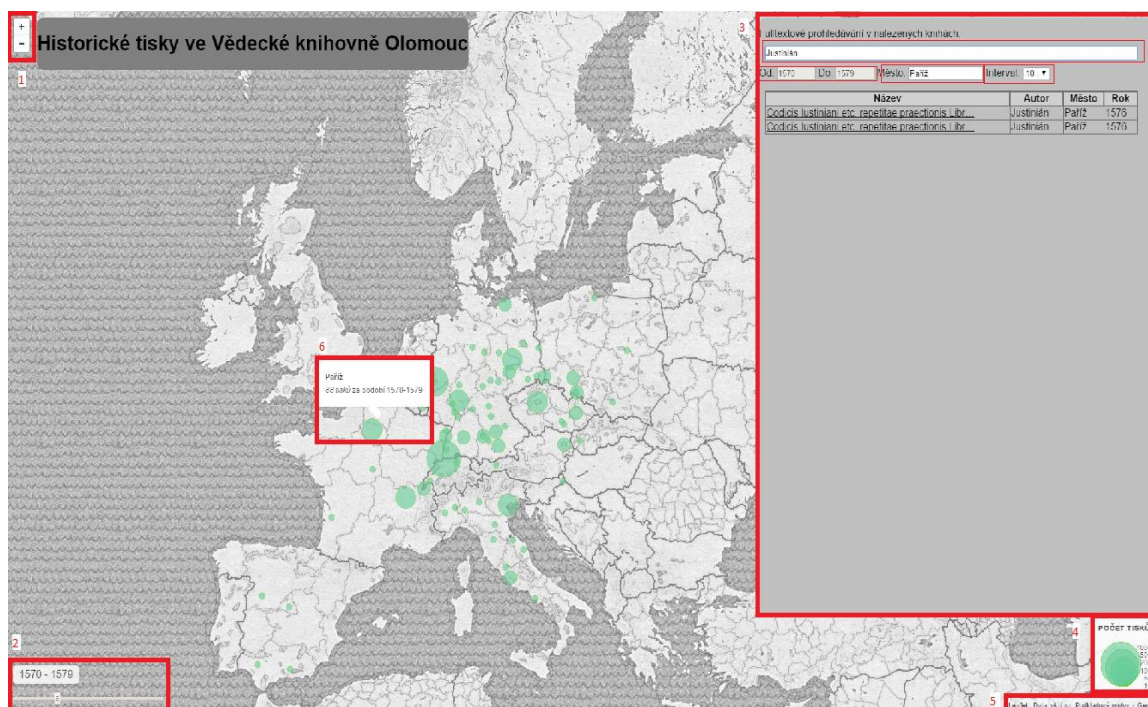
Pro dané účely bylo zvoleno využít geometrický bodový znak kruhového tvaru, který bude reprezentovat každé místo vydání. Podle počtu výtisků v daném období bude mít bodový znak různou velikost rozdělenou do kategorií, diskrétní stupnice je v legendě.

## 4 FUNKCIONALITA APLIKACE

Tato kapitola si klade za cíl popsat některé případy užití aplikace a funkcionalitu jednotlivých prvků.

### 4.1 Rozložení ovládacích prvků

Aplikace je vytvářena pomocí frameworku Leaflet, který poskytuje základní prohlížečku rastrových a vektorových digitálních map s volitelnými funkcemi. Bylo použito frameworku Leaflet[4] a bylo aplikováno řešení časové mapy podle tutoriálu „Time Series Proportional Symbol Maps with Leaflet and jQuery.“ [5]



Obr. 4 Screenshot z aplikace – rozložení základních prvků.

Uživatel je po návštěvě internetové stránky informován o hlavním obsahu díky názvu mapy umístěném vlevo nahoře. Další prvky jsou označeny na obrázku níže pomocí číslic, umístění a význam jednotlivých ovládacích prvků je popsán níže:

1. Vlevo nahoře je ovládací prvek pro **zoomování**,
2. Vlevo dole je **časový posuvník**, kterým se vybírá období, které chceme zobrazit.

3. Vpravo nahoře se nachází **tabulka s obsahem** zobrazující knihy z vybraného města a ovládací prvky pro další zúžení výběru, ty jsou zobrazeny lépe na obrázku níže
  - a. Pole pro fulltextové vyhledávání
  - b. Časové období (od, do) – uzamčené pro úpravy
  - c. Město – možnost ručně zadat jiné město (ruší výběr z mapy)
  - d. Interval – změna délky časového intervalu (roky)
4. Legenda
5. Tiráž
6. Aktivní prvek – zobrazený popisek s názvem města, časovým obdobím a počtem knih

Fulltextové prohledávání v nalezených knihách:

Justinián

Od: 1570 Do: 1579 Město: Paříž Interval: 10 ▼

Název	Autor	Město	Rok
Codicis Iustiniani etc. repetitae praefationis Libr...	Justinián	Paříž	1576
Codicis Iustiniani etc. repetitae praefationis Libr...	Justinián	Paříž	1576

Obr. 5 Detail zobrazení záznamů z konkrétního města a období s využitím fulltextového prohledávání v nalezených záznamech.

#### Zoomování:

Jedna ze základních funkcí, která je implementována je možnost přibližovat a oddalovat mapu a měnit tak měřítko mapy. Tuto funkcionalitu lze používat pomocí tlačítek

„+“ a „-“ v levém horním rohu, nebo pomocí kolečka myši. Jednorázové přiblížení je možné též udělat dvojklikem levým tlačítkem myši.

#### Časový posuvník:

Hlavním prvkem, který umožňuje ovládání dynamického aspektu aplikace je časový posuvník, zde lze měnit zvolený časový interval. Na základě změny období se okamžitě změní i obsah mapy.

#### Tabulka s obsahem:

V bočním menu, kde se nachází tabulka s obsahem se zobrazuje výčet záznamů ze zvoleného města. Je možné výběr zúžit zadáním nějakého hesla do fulltextového vyhledávače. Dále je možné ručně vypsát jiný název města (čímž se ale ruší výběr grafický). Je možné změnit interval – nabízené hodnoty jsou 1, 5, 10, 20, 50, 100. Po změně je nutné počkat až se data přepočítají a vykreslí se mapa znovu. Nové načtení vrátí časový posuvník na začátek, proto s ním je poté nutné pohnout do pozice, kdy jsou v mapě zobrazeny nějaké prvky.

Legenda:

Legenda popisuje velikosti jednotlivých symbolů v hraničních hodnotách.

Tiráž:

Tiráž popisuje základní informace o původu dat a autorovi.

Aktivní prvek:

Je zvolené město, vybrané kliknutím uživatele. Po kliknutí vyskočí okno s počtem tisků, názvem města a zvoleným obdobím. Na základě kliknutí na aktivní prvek se změní údaje v tabulce s obsahem.

## 5 VÝSLEDKY PRÁCE

Je vytvořena aplikace, která zobrazuje dostupná data z metadat o historických fondech Vědecké knihovny Olomouc. Aplikace zobrazuje jednotlivé záznamy agregované podle místa a roku vydání (resp. uživatelem zvoleného období).

Uživatel může kliknout na libovolné město a zobrazit tak detailně všechny tisky, které byly v daném městě ve vybraném období vydány. Z těchto knih může vybírat dále pomocí fulltextového vyhledavače.

Díky využití moderních principů tvorby webových aplikací (JavaScript, JQuery, AJAX) je aplikace poměrně plynulá.

Práce se zabývá stažením a zpracováním dat tak, aby byly využitelné pro potřeby aplikace. Nejvíce se potýká s nízkou kvalitou a vysokou chybovostí dat danou různými faktory, tuto chybovost do určité míry eliminuje.

## 6 DISKUZE

Aplikace nezobrazuje všechny položky z historických fondů. Bylo snahou autora práce zobrazit co nejvíce možných záznamů ze seznamu všech knih. Aby mohl záznam být v mapové aplikaci zobrazen, musí mít korektní informace o své poloze a roku vydání. Tyto informace žel u mnoha záznamů nebyly dostupné a autor v práci popisuje řešení, díky kterým se podařilo do značné míry zvýšit počet zobrazených záznamů v mapě.

V práci se vyskytují některé další chyby, např. vinou chybného geokódování sídel. Tyto problémy jsou autoru práce známy. Vzhledem ke špatné kvalitě informací o místu vydání, kvantitě dat, a časté absenci údajů o zemi vydání je téměř nemožné poskytnout stoprocentní záruku, že některé knihy byly vydány právě v tom místě, které mapa zobrazuje. Nicméně počet takových míst je zanedbatelný.

Práce je pohodlně zobrazitelná na počítačích s větší zobrazovací plochou, její responzivita však nebyla zajištěna, na mobilních zařízeních, kde nelze zobrazit celou mapu ani posouvat s posuvníkem pro změnu období.

Jednou z otázek ohledně výsledků práce může být i její užitečnost, nebo pozdější využitelnost v praxi. Jelikož jsou zpracovávána data historická a mají v sobě též prostorový aspekt, pravděpodobně si aplikace může najít uživatele mezi historiky a geografy. Badatelé zabývající se starými písemnostmi mohou mít lepší představu o geografickém rozložení míst, kde se vydávaly knihy a intenzitě produkce knih během různých období historie.

Webová mapa pravděpodobně neosloví širokou veřejnost. Protože však dosud kromě vyhledávacího katalogu neexistovaly jiné náhledy na databázi historických tisků ve Vědecké knihovně v Olomouci, může být tato práce inspirací pro využívání dostupných veřejných datových skladů a snahu o jejich další vizualizaci a získávání nových pohledů na již existující informace.

## **7 ZÁVĚR**

Cílem práce bylo zpracovat data o Historických tiscích Vědecké knihovny Olomouc a vytvořit jejich atraktivní prezentaci pomocí mapové aplikace zobrazující jejich časový a prostorový aspekt.

Byla vybrána vhodná řešení pro získání a vizualizaci dat. Kvalita vstupních dat, která znemožňovala u mnohých záznamů jejich vizualizaci byla částečně vylepšena pomocí postupů popsaných v práci. Chyby byly částečně eliminovány.

Aplikace atraktivním způsobem vizualizuje obsah Historických fondů ve Vědecké knihovně Olomouc. Metody popsané v práci je možné použít pro účely vizualizace v podobných projektech v budoucnu.

Aplikace najde využití u odborníků i laiků z oblasti geografie a historie, rovněž může posloužit návštěvníkům Vědecké knihovny Olomouc, kteří tím získávají nový pohled na knihy uložené v archivech knihovny.



## **SUMMARY**

The object of this diploma thesis was to create dynamic web map of Historical prints in Research Library Olomouc. Thesis contains brief introduction into theme, description of used methodology and technologies, description of functionality and tools used in the application.

Application is developed in PHP, uses AJAX, JQuery and JavaScript. Author had to deal with abnormalities and mistakes in data. For these problems were developed solutions. Application can be used by experts or lay workers in the field of geography, history or visitors of Research Library Olomouc.

## POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

- [1] KRAAK, M a Allan BROWN. Web cartography: developments and prospects. New York: Taylor, 2001, ix, 213 p. ISBN 07-484-0869-X.
- [2] Stylesheets for clean geographic data visualization [online]. 2017 [cit. 2017-05-02]. Dostupné z: <https://github.com/geoiq/acetate>
- [3] Mapbox [online]. 2017 [cit. 2017-05-02]. Dostupné z: <https://www.mapbox.com/>
- [4] Leaflet - a JavaScript library for interactive maps [online]. 2017 [cit. 2017-04-12]. Dostupné z: <http://leafletjs.com/>
- [5] DONOHUE, Richard G., Carl M. SACK a Robert E. ROTH. *Time Series Proportional Symbol Maps with Leaflet and jQuery* [online]. 2013 [cit. 2017-05-12]. Dostupné z: <http://www.cartographicperspectives.org/index.php/journal/article/view/cp76-donohue-et-al/1307>
- [6] MARC 21 Format for Bibliographic Data: Summary Statement of Content Designators (Network Development and MARC Standards Office, Library of Congress) [online]. 2016 [cit. 2017-04-12]. Dostupné z: <https://www.loc.gov/marc/bibliographic/bdsummary.html>
- [7] Vědecká knihovna v Olomouci [online]. 2017 [cit. 2017-03-05]. Dostupné z: <http://www.vkol.cz/cs/historicke-fondy/>
- [8] UNGEGN Working Group on Exonyms > Home [online]. 2017 [cit. 2017-03-05]. Dostupné z: <http://ungegn.zrc-sazu.si/>
- [9] W3Schools Online web tutorial [online]. 2017 [cit. 2017-03-05]. Dostupné z: <https://www.w3schools.com/>
- [10] Aleph.vkol.cz/OAI [online]. 2017 [cit. 2017-03-24]. Dostupné z: <http://aleph.vkol.cz/OAI>
- [11] Open archives initiative [online]. 2017 [cit. 2017-03-24]. Dostupné z: <https://www.openarchives.org/>
- [12] HANOUSEK, Tomáš. OAI-PMH pro začátečníky. Praha, 2007. Příloha č. 4 k výzkumné zprávě projektu VE20072009004. Dostupný z: [http://www.nacr.cz/wp-content/uploads/2015/11/moznosti\\_06.pdf](http://www.nacr.cz/wp-content/uploads/2015/11/moznosti_06.pdf).
- [13] Best XML Parser for PHP [duplicate]. *StackOverflow* [online]. 2017 [cit. 2017-02-10]. Dostupné z: <http://stackoverflow.com/questions/188414/best-xml-parser-for-php/>
- [14] OAI-PMH-PHP-Client/OAIPMHClient.php at master · kraenhansen/OAI-PMH-PHP-Client. *GitHub* [online]. 2017, 2017 [cit. 2017-02-10]. Dostupné z: <https://github.com/kraenhansen/OAI-PMH-PHP-Client/blob/master/src/oaipmh/OAIPMHClient.php>

## **PŘÍLOHY**

# SEZNAM PŘÍLOH

## Volné přílohy:

- Příloha 1 CD
- Příloha 2 Poster

## Popis struktury CD

Adresář css – css styly

Adresář import\_export – adresář s informacemi pro instalaci serveru a databáze

Adresář js – javascriptové knihovny

Adresář ostatní soubory – pracovní soubory Excel a Python skripty použité při tvorbě práce

Adresář web\_o\_webu – web o DP

diplomova\_prace\_haska.pdf – text samotné diplomové práce

ostatní soubory ve složce náleží funkcionalitě webu a aplikace