



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA INFORMAČNÍCH TECHNOLOGIÍ**

FACULTY OF INFORMATION TECHNOLOGY

**ÚSTAV INTELIGENTNÍCH SYSTÉMŮ**

DEPARTMENT OF INTELLIGENT SYSTEMS

**WEB PRO ZOBRAZOVÁNÍ ARCHIVÁLIÍ**

WEB FOR DISPLAYING ARCHIVE MATERIALS

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

**DAVID TŘÍSKA**

**VEDOUcí PRÁCE**

SUPERVISOR

**JAROSLAV ROZMAN, Ing., Ph.D.**

BRNO 2021

## Abstrakt

Cílem této bakalářské práce je průzkum stavu digitalizace archivů v České republice. Návrh a implementace vlastního webu pro zobrazování archivních dat, se scripty pro automatizované získávání těchto dat.

## Abstract

The aim of this bachelor thesis is to survey the state of digitization of archives in the Czech Republic. Design and implementation of own website for displaying archive data, with scripts for automatic acquisition of this data

## Klíčová slova

archiv, archiválie, web, PHP, Symfony, IIP, získávání dat z webu, obrázky s vysokým rozlišením

## Keywords

archives, archive record, web, PHP, Symfony, web scraping, high resolution images

## Citace

TRŽÍSKA, David. *Web pro zobrazování archiválií*. Brno, 2021. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Jaroslav Rozman, Ing., Ph.D.

# Web pro zobrazování archiválií

## Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana Ing. Jaroslava Rozmana, Ph.D.. Uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....  
David Tříška  
11. května 2022

## Poděkování

Rád bych poděkoval vedoucímu mé práce panu Ing. Jaroslavovi Rozmanovi, Ph.D. za veškeré poskytnuté rady a vedení.

# Zadání bakalářské práce



24661

Student: **Tříška David**  
Program: Informační technologie  
Název: **Web pro zobrazování archiválií**  
**Web for Displaying Archive Materials**  
Kategorie: Web

**Zadání:**

1. Nastudujte funkčnost a vzhled webů různých archivů v České republice. Zaměřte se na části, které zobrazují skeny archiválií.
2. Navrhněte web, který umožní zobrazovat požadované archiválie. Na základě rešerše archivů umožněte uživatelům co nejpohodlnější procházení scanů (stejně zvětšení obrázku při přechodu na další scan, ovládání klávesami, atd.). Web by měl být také snadno rozšiřitelný o další kategorie archiválií. Proto navrhněte skripty pro automatickou aktualizaci vybraných informací z archivů v České republice.
3. Podle návrhu web vytvořte a proveďte jeho důkladné otestování na skupině uživatelů.

**Literatura:**

- Matriční rozcestník, Česká genealogická a heraldická společnost v Praze, <http://www.genealogie.cz/aktivity/digitalizace/>, [cit. 15.10.2020]

Pro udělení zápočtu za první semestr je požadováno:

- První dva body zadání

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Rozman Jaroslav, Ing., Ph.D.**

Vedoucí ústavu: Hanáček Petr, doc. Dr. Ing.

Datum zadání: 1. listopadu 2021

Datum odevzdání: 11. května 2022

Datum schválení: 3. listopadu 2021



# Obsah

<b>1</b>	<b>Úvod</b>	<b>4</b>
<b>2</b>	<b>Průzkum současného stavu digitalizace archivů v České republice</b>	<b>5</b>
2.1	Archiv hlavního města Prahy . . . . .	6
2.2	Státní oblastní archiv v Litoměřicích . . . . .	7
2.3	Zemský archiv v Opavě . . . . .	7
2.4	Státní oblastní archiv v Zámrsku . . . . .	8
2.5	Státní oblastní archiv v Praze . . . . .	9
2.6	Státní oblastní archiv v Plzni . . . . .	11
2.7	Moravský zemský archiv v Brně . . . . .	12
2.8	Archiv města Ostravy . . . . .	13
2.9	Archiv Univerzity Karlovy v Praze . . . . .	13
2.10	Národní archiv České republiky . . . . .	14
2.11	Státní oblastní archiv v Třeboni . . . . .	15
<b>3</b>	<b>Vstupní data pro automatizované získávání obsahu</b>	<b>16</b>
<b>4</b>	<b>Zobrazování obrázků ve vysokém rozlišení na webu</b>	<b>23</b>
4.1	Obrazové formáty s více rozlišeními . . . . .	23
4.1.1	Tiled Multi-Resolution TIFF . . . . .	23
4.1.2	JPEG2000 . . . . .	24
4.2	Protokoly a aplikace . . . . .	24
4.2.1	IIPImage . . . . .	24
4.2.2	Zoomify . . . . .	24
4.2.3	OpenSeadragon . . . . .	25
<b>5</b>	<b>Návrh webu</b>	<b>26</b>
5.1	Návrh uživatelského rozhraní . . . . .	26
5.2	Návrh databáze . . . . .	29
5.3	Automatizované získávání obsahu . . . . .	29
5.4	Návrh prohlížeče archiválií . . . . .	30
<b>6</b>	<b>Implementace</b>	<b>31</b>
6.1	Použité technologie . . . . .	31
6.2	Serverová část . . . . .	32
6.3	Klientská část . . . . .	35
<b>7</b>	<b>Testování</b>	<b>37</b>

8 Závěr	38
Literatura	39

# Kapitola 1

## Úvod

V této práci se v první řadě zaměřuji na rešerši stavu digitalizace archivů v České republice. Na obsah jejich digitalizovaných záznamů a jejich přístupnost. Pozornost věnuji uživatelskému rozhraní jednotlivých archivů a technologiím, které využívají. V práci porovnávám jednotlivé archivy mezi sebou a hledám místo pro zlepšení.

Archivy obsahují úžasné množství informací, bohužel často se tyto informace nijak nemohou využít, protože k nim nemáme přístup. To se naštěstí v posledních letech mění a archivy zveřejňují čím dál tím více digitalizovaných záznamů. Pokud je prohlížení, vyhledávání a bádání pro uživatele nepřívětivé, nebo dokonce i otravné. Pomalu ztrácí celá digitalizace archivů smysl. To se právě v této práci snažím změnit. Cílem práce je vytvoření webové aplikace, která by uživatelům poskytla jednotné uživatelské rozhraní napříč archivy a sjednotila obsah pod jeden web, ze kterého badatel bude moci pohodlně prohlížet záznamy všech archivů.

Obsah práce je dělen do kapitol, ve kterých podrobně popisují jednotlivé kroky v tvorbě takového webu. Nejprve se tedy věnuji průzkumu a zhodnocení stavu jednotlivých archivů. Jelikož archivy obsahují mnoho různých kategorií, pro které je potřeba zvolit jiné způsoby zpracování, věnuji se průzkumu především matričních záznamů. Které budou první kategorií na tomto webu. Tomu se věnuji v první kapitole, průzkum současného stavu [2](#). Při tomto průzkumu jsem se rovnou zaměřil na data (odkazy), které budou sloužit jako vstupní argumenty scriptům dolujícím data z archivů, přehled těchto dat uvedených podle jednotlivých archivů je v kapitole [3](#).

Dalé je potřeba samotný web správně navrhnout, musí být v budoucnu snadno rošitelný o další kategorie, data vybraných kategorií je potřeba automatizovaně plnit a aktualizovat. Nesmím však ani zapomenout na kvalitní návrh uživatelského rozhraní, především tedy části pro zobrazování jednotlivých digitalizovaných záznamů. Toto všechno podrobně probírám v kapitole návrh webu [5](#).

Po průzkumu současného stavu a návrhu řešení je na čase se pustit do samotné implementace. V kapitole Implementace [6](#) Vás podrobně provedu mojí cestou a zkušeností s implementací tohoto webu. V této kapitole nejprve zaměřím na technologie, na kterých bude řešení postavené, poté se zaměříme na seznámení s architekturou celé aplikace a v poslední řadě s implementačními detaily jednotlivých částí.

Když už mám web naprogramovaný, mohlo by Vás napadnout, že mám hotovo. Opak je však pravdou, žádný systém není hotový dokud není řádně otestován. Testováním Vás provedu poslední kapitole [7](#)

## Kapitola 2

# Průzkum současného stavu digitalizace archivů v České republice

V této kapitole se věnuji průzkumu a zhodnocení aktuálního stavu jednotlivých archivů. Především částí, které jsou zaměřené na zobrazování skenů archiválií matričních záznamů.

U každého archivu věnuji pozornost především logické struktuře členění záznamů a jakým způsobem jsou generovány adresy (jednotlivých archiválií, jejich scanů, ale i členění sekcí). Další důležitou částí na kterou se v tomto průzkumu zaměřuji je uživatelské rozhraní. Jeho části, které slouží k prohlížení archiválií, přechodu mezi scany jednotlivých archiválií, příblížení a přechodu na další archiválii.

Seznam archivů, obsahujících digitalizované matriční záznamy, je uveden na webových stránkách České genealogické a heraldické společnosti v Praze [4].

- Archiv hlavního města Prahy [2.1](#)
- Státní oblastní archiv v Litomeřicích [2.2](#)
- Státní oblastní archiv v Zámrsku [2.4](#)
- Zemský archiv v Opavě [2.3](#)
- Státní oblastní archiv v Třeboni [2.11](#)
- Státní oblastní archiv v Praze [2.5](#)
- Státní oblastní archiv v Plzni [2.6](#)
- Moravský zemský archiv v Brně [2.7](#)
- Archiv města Ostravy [2.8](#)
- Archiv Univerzity Karlovy v Praze [2.9](#)
- Národní archiv České republiky [2.10](#)

## 2.1 Archiv hlavního města Prahy

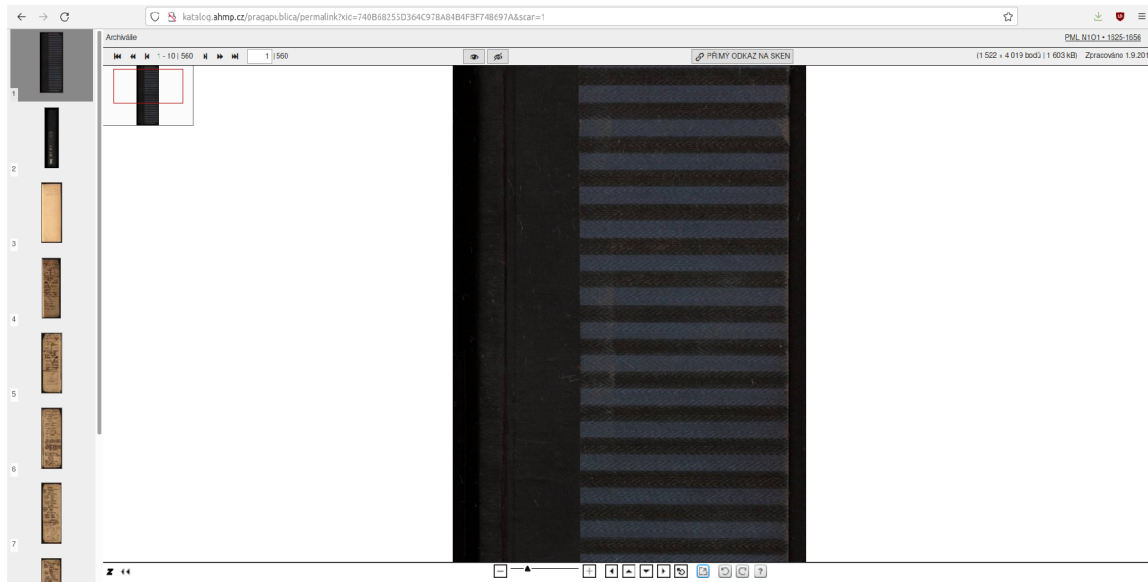
Archiv hlavního města Prahy nabízí badatelům dva přístupy pro vyhledávání v matričních záznamech. Využití archivního katalogu [1] nebo pomocí archivní pomůcky pro sbírku matrik [2]. Při vyhledávání pomocí archivního katalogu lze filtrovat záznamy pomocí jednotlivých sekcí nebo zobrazení všech matričních záznamů najednou. Při vyhledávání pomocí archivní pomůcky pro sbírku matrik je pro každou sekci (farnost, úřad) vygenerován PDF soubor se soupisem příslušných matrik.

Matriční záznamy arhivu hlavního města Prahy jsou členěny do následujících základních kategorií.

- matriky katolické
- matriky ostatních církví
- matriky civilní

Okno pro zobrazení archiválie je rozděleno do čtyřech částí. Horní ovládací lišta, spodní ovládací lišta, postranní levý panel zobrazující náhledy dalších stránek archiválie a hlavní okno zobrazující aktuálně prohlížený scan.

Uživatelské rozhraní pro zobrazování archiválií je zobrazeno na obrázku 2.1. Pro zobrazování archiválií je využita aplikace *Zoomify* 4.2.2. Jednotlivé scany záznamu je možné přibližovat a oddalovat tyto operace probíhají hladce a bez přílišných časových prodlev. Posunování zobrazované části lze provádět pomocí myši i klávesnice. Lze využít uzamknutí zvětšení a pozice zobrazované části. Při přechodu mezi jednotlivými scany se poté zobrazuje požadovaná oblast v daném zvětšení. Přechod na další scan lze provádět pouze myší, kliknutím na miniaturu požadovaného scanu.



Obrázek 2.1: Uživatelské rozhraní pro zobrazení archiválie, Archiv hlavního města Prahy. První scan archiválie PML N101 • 1625-1656<sup>1</sup>.

<sup>1</sup><http://katalog.ahmp.cz/pragapublica/permalink?xid=740B68255D364C978A84B4FBF748697A&scan=1>

## 2.2 Státní oblastní archiv v Litoměřicích

Státní oblastní archiv v Litoměřicích odkazuje ze své elektronické badatelný [12] přímo do databáze matrik [11]. V této sekci matriky nejsou děleny do specifických sekcí, lze ovšem využít vyhledávání a řazení.

Uživatelské rozhraní pro zobrazování archiválií je zobrazeno na obrázku 2.2. Pro zobrazování archiválií je použita aplikace *Zoomify* 4.2.2. Rozhraní je téměř totožné s archivem hlavního města Prahy 2.1. Hlavním rozdílem je zobrazení 25 miniatur snímků. Stejně jako u Archivu hlavního města Prahy 2.1 při přechodu na poslední scan na stránce, se nenčte další strana miniatur. Tuto akci je nutné vyvolat tlačítkem *přechod na další stranu* v horní ovládací liště.



Obrázek 2.2: Uživatelské rozhraní pro zobrazení archiválie, Státní oblastní archiv v Litoměřicích. První scan archiválie O • sig. 146/21 • 1881 - 1926 • Krásný Les<sup>2</sup>.

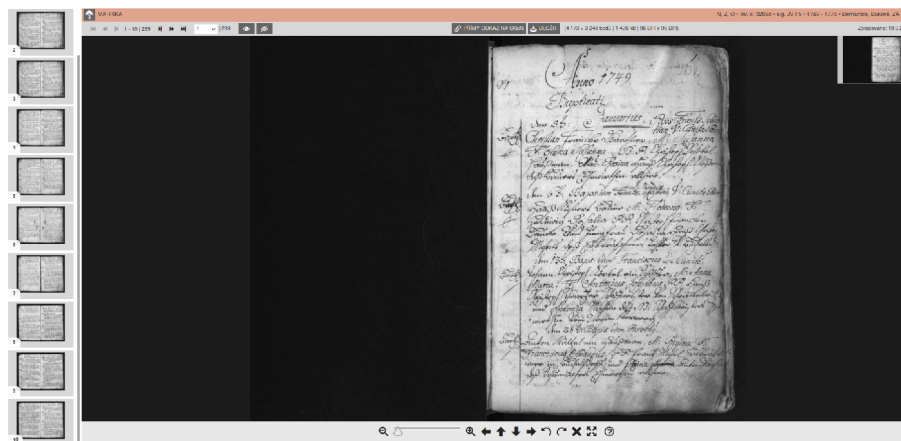
## 2.3 Zemský archiv v Opavě

Zemský archiv v Opavě nabízí ze svých webových stránek přímý odkaz do databáze digitalizovaných matričních záznamů[5]. Uživatelské rozhraní i logické členění je téměř totožné se Státním oblastním archivem v Litoměřicích 2.2. Matriční databáze není členěna např. dle úřadů, obsahuje ovšem vyhledávač. Řazení vyhledaných výsledků je možné dle stejných kategorií, jako u Litoměřického archivu 2.2. Matriční záznamy jsou opatřeny inventárním číslem, signaturou, obsahem a územním rozsahem.

Uživatelské rozhraní pro zobrazování snímků archiválií je na obrázku 2.3. K zobrazování scanů matričních záznamů je využita JavaScriptová aplikace *Zoomify* 4.2.2. Rozhraní pro prohlížení scanů je podobné s archivy 2.1, 2.2.

<sup>2</sup><http://vademecum.soalitomeric.cz/vademecum/permalink?xid=09ddd7cea03b9b8d30bdd2c7:1201ea2ef5b:-7e0b&scan=d4cf69fce28148149cfd43f3676478a>

<sup>3</sup><https://digi.archives.cz/da/permalink?xid=be89b34c-f13c-102f-8255-0050568c0263&scan=f0c2acd4516c42e0872a0d1122e81f2c>



Obrázek 2.3: Uživatelské rozhraní pro zobrazení archiválie, Zemský archiv v Opavě. První scan archiválie N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková<sup>3</sup>.

- [databáze matričních záznamů](#)
- [url pro přechod na další strany databáze matričních záznamů](#)
  - pro přístup potřeba platná cookie
- [archiválie \(např. N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková\)](#)
- [url prvního scanu archiválie \(např. první scan N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková\)](#)
  - pro přístup potřeba platná cookie
- [url pro procházení stránek scanů archiválie](#)
  - pro přístup potřeba platná cookie
- [permalink scanu archiválie \(např. prvního scanu N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková\)](#)
- [zoomify složka \(např. prvního scanu N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková\)](#)

## 2.4 Státní oblastní archiv v Zámrsku

Státní oblastní archiv v Zámrsku přechází na nový informační systém, který by měl být v provozu v roce 2022. V současné době archiv nenabízí žádný vlastní nástroj pro prohlížení digitalizovaných archiválií. Přístup k digitalizovaným matričním záznamům je umožněn přes inventurní seznam matričních záznamů<sup>4</sup>. Tento seznam je PDF soubor, aktualizovaný naposledy 18.8.2020. Obsahuje matriční záznamy děleny dle úřadů, které matriky vedly a územního obvodu. Každý záznam je tvořen číslem signatury, obsahem matriky (O - matrika oddaných, N - matrika narozených, Z - matrika zemřelých) a odkazem na stažení

<sup>4</sup>[https://stare.vychodoceskearchivy.cz/zamrsk/files/2020/08/8700\\_Sbrika-matrik-Vychodoceskeho-kraje-1587-1949\\_NAD\\_190-stav-2020-08-18.pdf](https://stare.vychodoceskearchivy.cz/zamrsk/files/2020/08/8700_Sbrika-matrik-Vychodoceskeho-kraje-1587-1949_NAD_190-stav-2020-08-18.pdf)



zip souboru, obsahující digitalizované scany. Dělení úřadů, které vedly matriční záznamy je následující.

- Římskokatolické církve
- Evangelické církve
  - Reformované a luterské sbory a kazatelské stanice (Českobratrská církev evangelická, Německá evangelická církev)
  - Obnovená jednota bratrská (ochranovská)
  - Svobodná česká evangelická církev (tzv. balcariáni)
  - Církev československá
- Civilní (okresní) úřady

## 2.5 Státní oblastní archiv v Praze

Státní oblastní archiv v Praze nabízí ze své eBadatelny<sup>[13]</sup> odkaz přímo do matričního vyhledávače<sup>[14]</sup>. V matričním vyhledávači lze vyhledávat záznamy podle klíčového slova, názvu matričních záznamů nebo lokality. Dalším způsobem vyhledávání je podle typu průvodce, s výběrem určité lokality. Typem průvodce je úřad.

- Římskokatolická církev
- Evangelická církev
- Československá církev
- Okresní úřad
- Jednota bratrská
- Stavovský úřad

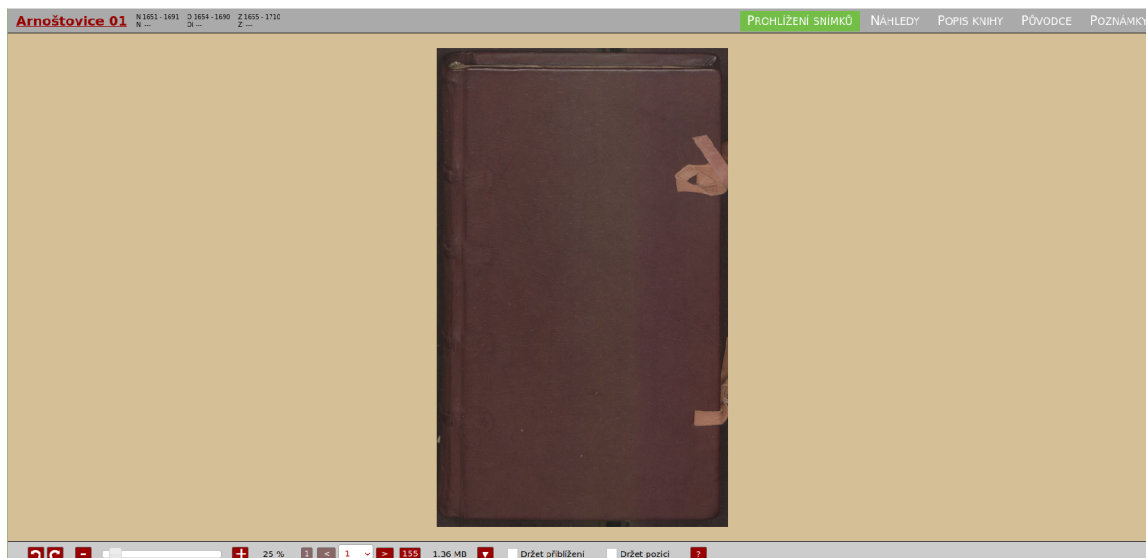
Vyhledané výsledky lze omezit časovým rozmezím. Vyhledávání podle časového období, lze použít i bez předchozích filtrů, tento způsob vyhledávání je jednodušší cestou pro získání všech matričních záznamů. Časový filter je ve výchozí hodnotě nastaven od roku 1000 do roku 2100, tímto vyhledáním, lze získat všechny matriční záznamy, rozdělené na strany.

Pro zobrazování jednotlivých scanů archiválií není využita žádná aplikace třetí strany na straně klienta nebo serveru. Zobrazovat snímky archiválie, lze po jednom scanu, obrázek 2.4 nebo jako náhledy po více scanech na stránce, obrázek 2.5. U náhledů je možnost volby několika různých velikostí od 8 snímku na stránku, až po 126 snímků na stránku. Každý náhled slouží jako odkaz na scan v plné velikosti.

---

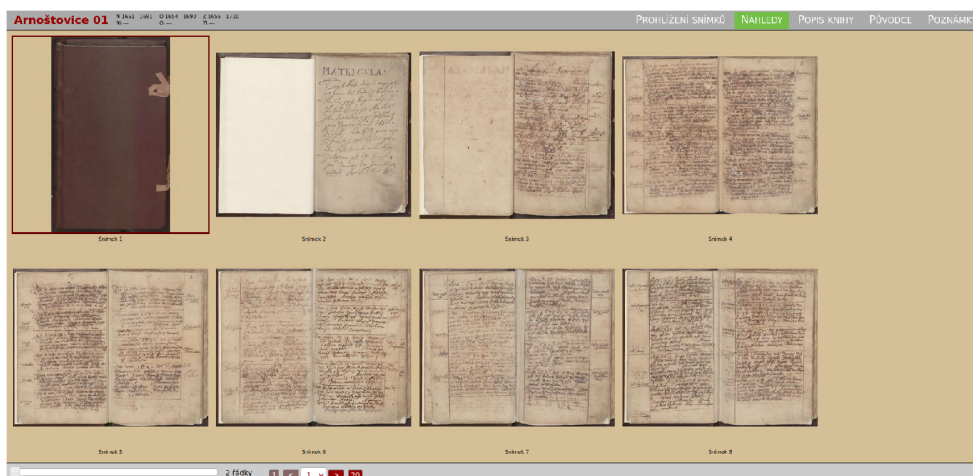
<sup>5</sup><https://ebadatelna.soapraha.cz/d/3726/1>





Obrázek 2.4: Uživatelské rozhraní pro zobrazení archiválie, Státní oblastní archiv v Praze. První scan archiválie Arnoštovice 01<sup>5</sup>.

Uživatelské rozhraní nabízí možnost přiblížení, oddálení, uzamknutí přiblížení i pozice při přechodu na další stránky. Ovládání je možné i pomocí kláves, ne ovšem všechny akce. Přiblížení lze provádět klávesami *shift* nebo *+* na numerické klávesnici, oddálení pomocí kláves *ctrl* nebo *-* na numerické klávesnici. Přechod mezi scany pomocí kláves *PgUp* nebo *PgDn*, není dle mého názoru příliš vhodný. Chybí možnost pohybu pomocí kláves zobrazovou částí při přiblížení.



Obrázek 2.5: Uživatelské rozhraní pro zobrazení archiválie, Státní oblastní archiv v Praze. Zobrazení náhledu archiválie Arnoštovice 01<sup>6</sup>.

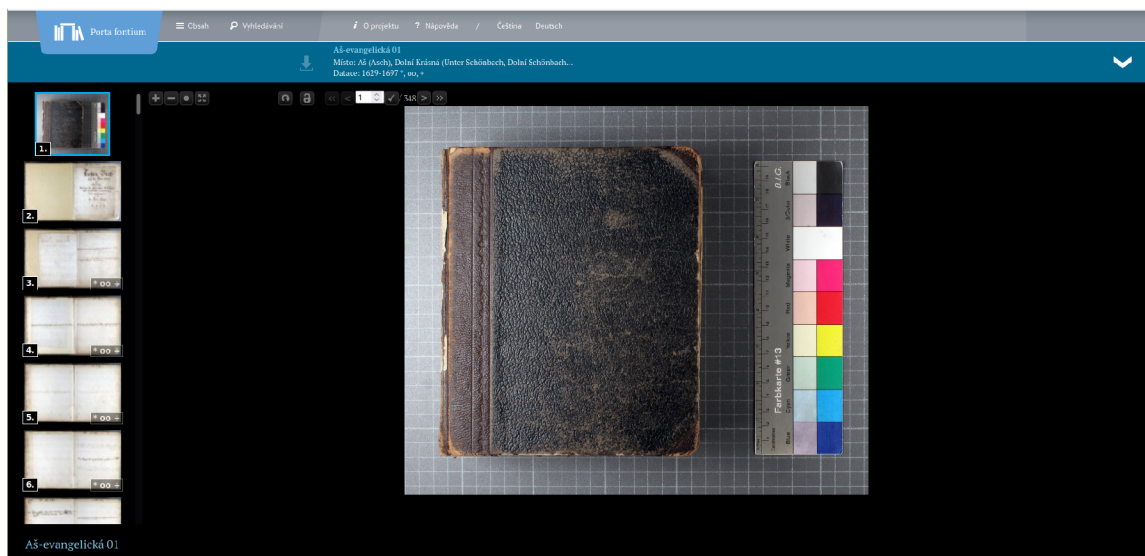
<sup>6</sup><https://ebadatelna.soapraha.cz/d/3726/1>

## 2.6 Státní oblastní archiv v Plzni

Státní oblastní archiv v Plzni zveřejňuje digitalizované matriční záznamy na webu Porta fontium [3], projekt spojující české a bavorské digitální historické prameny. Odsud vede odkaz do sekce *Sbírka matrik západních Čech*<sup>7</sup>, které jsou děleny následovně.

- Církev evangelická (1581-1949)
- Církev římskokatolická (1539-1949)
- Církev československá (1923-1949)
- Církev československá evangelická (1915-1949)
- Církev Jednota bratrská (1945-1949)
- Církev řeckokatolická (1946-1949)
- Církev pravoslavná (1948-1949)
- Okresní úřad, národní výbor (1869-1949)
- Standesamt (1938-1945) - matriky vedené německými civilními úřady

Káždá z těchto sekcí je dále dělena dle územního rozsahu. Stránky územního rozsahu obsahují konkrétní matriční záznamy a také matriční záznamy, které budou doplněny. Stránka archiválie obsahuje kromě popisu svého obsahu, také náhledy všech snímků archiválie, které slouží jako odkazy do prohlížeče jednotlivých snímků.



Obrázek 2.6: Uživatelské rozhraní pro zobrazení archiválie, Státní oblastní archiv v Plzni. První scan archiválie Aš-evangelická 01<sup>8</sup>.

<sup>7</sup><https://www.portafontium.eu/contents/register/>

<sup>8</sup>[https://www.portafontium.eu/iipimage/30060059/as-evangelicka-01\\_0010-x](https://www.portafontium.eu/iipimage/30060059/as-evangelicka-01_0010-x)

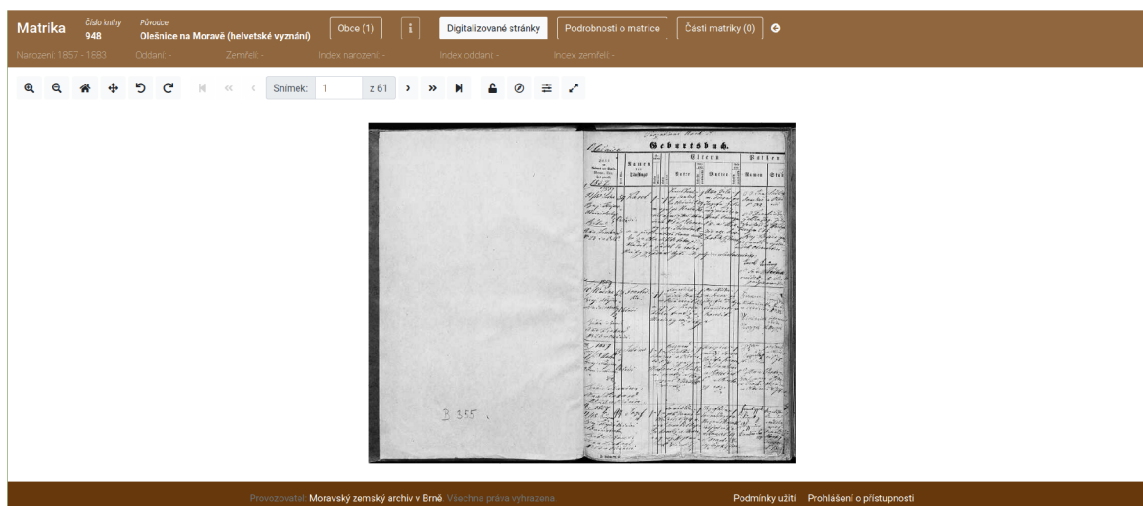
Pro prohlížení snímků je použita JavaScriptová aplikace *OpenSeaDragon*. Uživatelské rozhraní, na obrázku 2.6 je rozděleno do třech částí. Horní navigační lišta, obsahující tlačítka pro přiblížení, oddálení, rotaci snímku, uzamknutí nastavení při přechodu na další snímky, tlačítka pro pohyb mezi scany, výběr konkrétního scanu podle jeho pořadí, a zobrazení podrobností matričního záznamu. Další částí je levá postranní lišta zobrazující miniatury dalších scanů, které slouží jako odkazy na tyto scany. Jsou načteny miniatury všech scanů, takže přechod je plynulejší. Poslední a hlavní prvek je okno pro samotné zobrazení archiválie. Některé akce lze ovládat i pomocí klávesnice, jedná se o přiblížení, oddálení a pohyb v zobrazovné části.

## 2.7 Moravský zemský archiv v Brně

Matriční záznamy Moravského zemského archivu v Brně jsou přístupné v sekci *Actapublice*[9]. Tato sekce slouží jako vyhledávač matričních záznamů. Vyhledávat je možné podle obce, čísla knihy nebo podle původce. Vyhledávací filtry je možné kombinovat.

Výsledkem vyhledávání je seznam matričních záznamů, u každého záznamu jsou uvedeny následující položky číslo knihy, původce, obsah (narození, zemřelí, oddaní), obec a počet digitalizovaných snímků. Záznam slouží jako odkaz do prohlížeče scanů archiválie.

Pro prohlížení snímků je využita JavaScriptová aplikace *OpenSeaDragon*. Uživatelské rozhraní, na obrázku 2.7 je rozděleno do tří částí horizontální lišta zobrazující informace o matričním záznamu, horizontální ovládací lišta a hlavní okno zobrazující scan archiválie. Ovládací lišta obsahuje tlačítka pro přiblížení, oddálení, zobrazení přes celou obrazovku, přechod na výchozí pozici a zvětšení, přechod mezi snímky, uzamknutí nastavení při přechodu na další snímek a nastavení pro změnu kontrastu a jasů.



Obrázek 2.7: Uživatelské rozhraní pro zobrazení archiválie, Moravský zemský archiv v Brně. První scan archiválie č. knihy 948, Olešnice na Moravě (helvetské vyznání)<sup>9</sup>.

## 2.8 Archiv města Ostravy

Archiv města Ostravy uchovává pouze domovské matriční záznamy. Archiv ze své digitální badatelny<sup>[6]</sup> nabízí odkazy do nejnavštěvovnějších sekcí, jednou z těchto sekcí je i sekce *Domovské matriky*. Celá digitální badatelna je tvořena formou vyhledávače, uživatel zvolí v jakých sekcích vyhledávat, nabízené odkazy nejnavštěvovnějších sekcí konfiguruje filtry vyhledávače.

Pro zobrazování archiválií je použita JavaScriptová aplikace *Zoomify* 4.2.2. Uživatelské rozhraní je proto totožné s archivy Archiv hlavního města Prahy 2.1, Státní oblastní archiv v Litoměřicích 2.2, Zemský archiv v Opavě 2.3.

## 2.9 Archiv Univerzity Karlovy v Praze

Archiv Univerzity Karlovy v Praze obsahuje projekt *Studenti pražských univerzit 1882–1945* ve kterém digitalizuje matriční záznamy studentů univerzity. Matriční záznamy jsou rozděleny do dvou kategorií. Tyto sekce obsahují matriční záznamy dělené podle období.

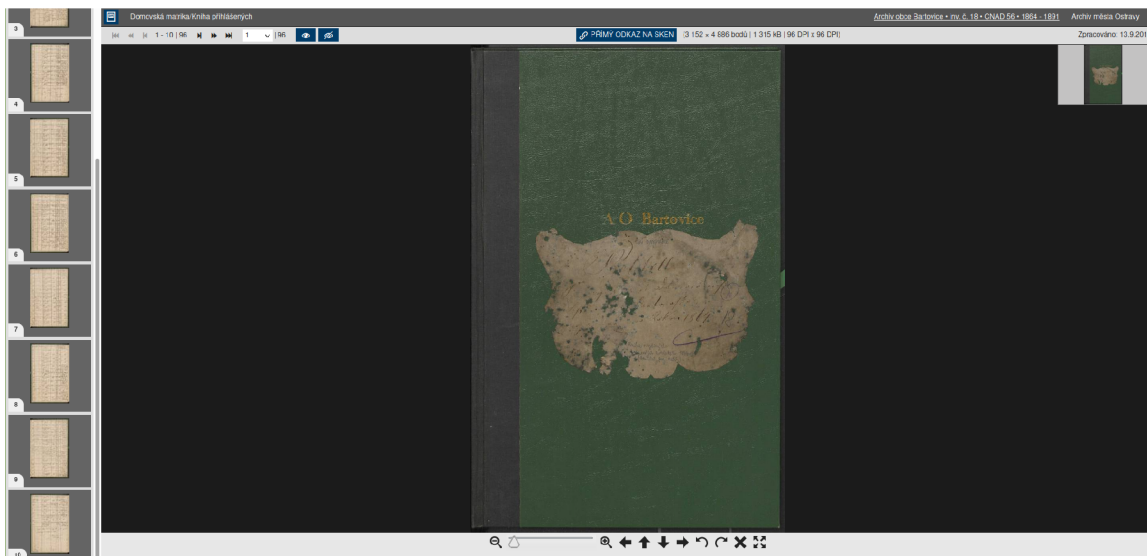
- Matriky Univerzity Karlovy
- Matriky Německé univerzity v Praze

Pro prohlížení scanů archiválií není využita žádná aplikace třetí strany. Uživatelské rozhraní, obrázek 2.9, není příliš přívětivé. Při prohlížení jsou zobrazeny informace o archiválii, následuje zobrazení aktuálního scanu, ve spodní části je navigační lišta pro přechod mezi snímky. Scan je možné přibližovat, oddalovat, otáčet. Ovládání není možné pomocí kláves.

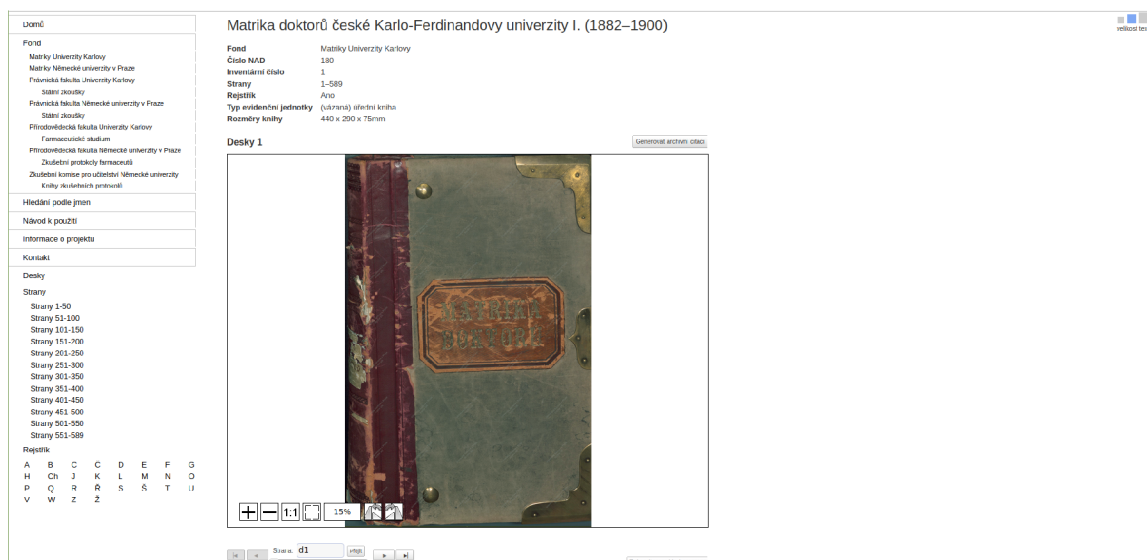
<sup>9</sup><https://www.mza.cz/actapublica/matrika/detail/3734?image=21600010-000253-003368-000000-000948-000000-00-B01996-00010.jp2>

<sup>10</sup><https://badatelna.ostrava.cz/vademecum/permalink?xid=CD9DB94AA15311E9B75C000C29D8B843&scan=06c131450a824acaa1f5f907d70b94b8>

<sup>11</sup><https://is.cuni.cz/webapps/archiv/public/book/bo/1889173198808193/1/?lang=cs>



Obrázek 2.8: Uživatelské rozhraní pro zobrazení archiválie, Archiv města Ostravy. První scan archiválie Archiv obce Bartovice • inv. č. 18 • CNAD 56 • 1864 - 1891<sup>10</sup>.



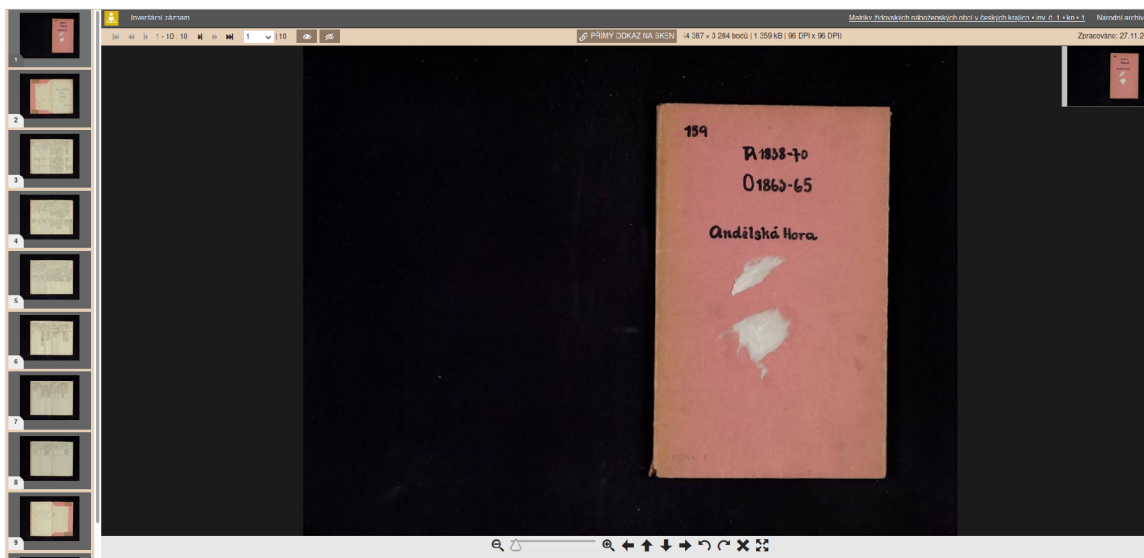
Obrázek 2.9: Uživatelské rozhraní pro zobrazení archiválie, Archiv Univerzity Karlovy v Praze. První scan archiválie Matrika doktorů české Karlo-Ferdinandovy univerzity I. (1882–1900)<sup>11</sup>.

## 2.10 Národní archiv České republiky

V současné době jsou v Národním archivu České republiky digitalizovány Židovské matriky z období 1784 - 1949. Přístup k těmto záznamům je možný přes archivní pomůcku *židovské matriky 1784 - 1949 (1960)* [10].

Uživatelské rozhraní, na obrázku 2.10, používá *Zoomify 4.2.2*. Rozhraní je obdobné jako u archivů 2.1, 2.2, 2.8.





Obrázek 2.10: Uživatelské rozhraní pro zobrazení archiválie, Archiv Univerzity Karlovy v Praze. První scan archiválie Andělská Hora (o. Karlovy Vary)<sup>12</sup>.

## 2.11 Státní oblastní archiv v Třeboni

Tento archiv uvádí, že jakékoli prohlížení snímků mimo jejich webové rozhraní je porušením autorských práv.

### Shrnutí

Z průzkumu jsem zjistil následující. Většina archivů využívá některou z aplikací pro zobrazování obrázků třetích stran. Jedná se především o aplikace *Zoomify* 4.2.2, kterou využívá celkem 5 archivů a aplikace *OpenSeaDragon* 4.2.3. Ostatní archivy zobrazují obrázky v plné velikosti. Z testování uživatelského rozhraní pro zobrazování archiválií jednotlivých archivů, jsem objevil společnou negativní vlastnost. Je to přechod na další scan archiválie. Při kterém dochází k přesměrování na novou stránku, jejíž načtení není zrovna rychlé. Mimo jiné se znovu stahují náhledy vybraných následujících a předcházejících snímků. Ty však zůstávají stejné v rámci jedné stránky (většinou obsahující 10 snímků). V tomto bych viděl možnost pro vylepšení.

Dalším negativním prvkem je obtížnost přechodu z domovské stránky archivu do digitálního archivu. Často chyběl odkaz do této sekce na hlavní stránce.

Při zobrazování archiválií většina archivů obsahovala možnost ovládání pomocí klaves. To jsem velmi vítal, bohužel často chyběl popis ovládání. Takže jsem musel zkoušet, ovládání bylo podobné napříč archivy, ale ne shodné. Proto bych uvítal nějaký popis ovládání.

<sup>12</sup><https://vademecum.nacr.cz/vademecum/permalink?xid=69f27506-24a8-41eb-890d-c8dccc62b4562&scan=28c986cdcea507bf98ae3bea4d807106>

## Kapitola 3

# Vstupní data pro automatizované získávání obsahu

Pokud chceme, aby aplikace automatizované získávala obsah z jednotlivých kategorií archivů musíme se podívat jakým způsobem jsou archivy dělené. Zaměřil jsem se na kategorii matričních záznamů. Při přidávání další sekce, bude potřeba provést podobný průzkum a implementovat skripty. Aktualizační skripty potřebují znát odkazy ze kterých mají data stahovat a jak se dostat na další stránky. U kategorie matriční záznamy jsou to tyto stránky, pro každý archiv se může částečně lišit.

- archivní katalog, seznam všech záznamů v kategorii
- záznam, stránka konkrétního záznamu
- první scan záznamu
- metadata obrázku

### Archiv hlavního města Prahy 2.1

- [archivní katalog](#)

```
http://katalog.ahmp.cz/pragapublica/permalink
?xid=7EF18906B65E11DF820F00166F1163D4
```

- [kategorie nejnižší úrovně \(např. Kostel P. Marie na Louži\)](#)

```
http://katalog.ahmp.cz/pragapublica/permalink
?xid=1e81f068b865ae876ffe0ce67f99d28e
```

- [archiválie \(např. PML N1O1 • 1625-1656\)](#)

```
http://katalog.ahmp.cz/pragapublica/permalink
?xid=740B68255D364C978A84B4FBF748697A
```

- [scan archiválie \(např. první scan PML N1O1 • 1625-1656\)](#)

[http://katalog.ahmp.cz/pragapublica/permalink  
?xid=740B68255D364C978A84B4FBF748697A  
&scan=1](http://katalog.ahmp.cz/pragapublica/permalink?xid=740B68255D364C978A84B4FBF748697A&scan=1)

- [zoomify složka \(např. prvního scanu PML N1O1 • 1625-1656\)](#)

[http://images.ahmp.cz/mrimage/ahmp  
/zoomify/cz/archives/CZ-321100010  
/NAD-156/dao/images/0060  
/2a840aac78fbe900119a79e611f5b467.jpg](http://images.ahmp.cz/mrimage/ahmp/zoomify/cz/archives/CZ-321100010/NAD-156/dao/images/0060/2a840aac78fbe900119a79e611f5b467.jpg)

## Státní oblastní archiv v Litoměřicích 2.2

- [databáze matričních záznamů](#)

[http://vademecum.soalitomeric.cz  
/vademecum/searchlink  
?fcAr214000010-0  
&fcDb=10041](http://vademecum.soalitomeric.cz/vademecum/searchlink?fcAr214000010-0&fcDb=10041)

- [url pro přechod na další strany databáze matričních záznamů](#)

– pro přístup potřeba platná cookie

[http://vademecum.soalitomeric.cz  
/vademecum/PaginatorResult.action  
?\\_sourcePage=-ve93-Zy9f8JJBf9VWwgMmyThV5RMAPX\\_6BJHewyPBiqxLxB8  
WDp1gqhuK44y90EYaSI958xXxCAGTLAvSyLU404LLfhJZ4XzFuiQVKK-k\%3D  
&row=12](http://vademecum.soalitomeric.cz/vademecum/PaginatorResult.action?_sourcePage=-ve93-Zy9f8JJBf9VWwgMmyThV5RMAPX_6BJHewyPBiqxLxB8WDp1gqhuK44y90EYaSI958xXxCAGTLAvSyLU404LLfhJZ4XzFuiQVKK-k\%3D&row=12)

- [archiválie \(např. O • sig. 146/21 • 1881 - 1926 • Krásný Les\)](#)

[http://vademecum.soalitomeric.cz  
/vademecum/permalink  
?xid=09ddd7cea03b9b8d:30bdd2c7:1201ea2ef5b:-7e0b](http://vademecum.soalitomeric.cz/vademecum/permalink?xid=09ddd7cea03b9b8d:30bdd2c7:1201ea2ef5b:-7e0b)

- [první scan archiválie \(např. první scan O • sig. 146/21 • 1881 - 1926 • Krásný Les\)](#)

– pro přístup potřeba platná cookie

[http://vademecum.soalitomeric.cz  
/vademecum/Zoomify.action?entityRef=  
\%28\%5En\%29\%28\%28\%28localArchiv\%2C\%5En\%28\%29\%28unidata\%29\%29\%28137646\%29\%29  
&scanIndex=0](http://vademecum.soalitomeric.cz/vademecum/Zoomify.action?entityRef=%28\%5En\%29\%28\%28\%28localArchiv\%2C\%5En\%28\%29\%28unidata\%29\%29\%28137646\%29\%29&scanIndex=0)

- [procházení stránek scanů archiválie](#)

– pro přístup potřeba platná cookie



[http://vademecum.soalitomerice.cz/vademecum/PaginatorMedia.action?\\_sourcePage=q-1-wB2AMjaUGokIEH7Q9d9-XjRTFBsNSEhS-qNkcSXjoppFd17G\\_z0UgHzBODPB1lua1kc-pbq9\\_CHpyX2B\\_YvEuSTM12C1tq2Ecr2VZk\%3D&row=25](http://vademecum.soalitomerice.cz/vademecum/PaginatorMedia.action?_sourcePage=q-1-wB2AMjaUGokIEH7Q9d9-XjRTFBsNSEhS-qNkcSXjoppFd17G_z0UgHzBODPB1lua1kc-pbq9_CHpyX2B_YvEuSTM12C1tq2Ecr2VZk\%3D&row=25)

- [permalink scanu archiválie \(např. prvního scanu O • sig. 146/21 • 1881 - 1926 • Krásný Les\)](#)

<http://vademecum.soalitomerice.cz/vademecum/permalink?xid=09ddd7cea03b9b8d:30bdd2c7:1201ea2ef5b:-7e0b&scan=d4cf69fce28148149cfdd43f3676478a>

- [zoomify složka \(např. prvního scanu O • sig. 146/21 • 1881 - 1926 • Krásný Les\)](#)

<http://images.soalitomerice.cz/mrimage/matriky/zoomify/cz/archives/CZ-214000010/NAD-856/dao/images/0117/695eaa7d-3121-45a8-8457-72fec7290614.jpg>

## Zemský archiv v Opavě 2.3

- [databáze matričních záznamů](#)

<https://digi.archives.cz/da/searchlink?fcDb=10041>

- [url pro přechod na další strany databáze matričních záznamů](#)

– pro přístup potřeba platná cookie

[https://digi.archives.cz/da/PaginatorResult.action?\\_sourcePage=yZsJgT604pM2ENBbJUtfYhJxs0HAPrMBPdNL5sGFBSkPwv70xVk\\_EP7EJ-iYh\\_YGot8-1YFlstLP8e\\_jI2XqMkfgHcUsLQtX09YwDD0ypeI\%3D&row=12](https://digi.archives.cz/da/PaginatorResult.action?_sourcePage=yZsJgT604pM2ENBbJUtfYhJxs0HAPrMBPdNL5sGFBSkPwv70xVk_EP7EJ-iYh_YGot8-1YFlstLP8e_jI2XqMkfgHcUsLQtX09YwDD0ypeI\%3D&row=12)

- [archiválie \(např. N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková\)](#)

<https://digi.archives.cz/da/permalink?xid=be89b34c-f13c-102f-8255-0050568c0263>

- [url prvního scanu archiválie \(např. první scan N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková\)](#)

– pro přístup potřeba platná cookie

[https://digi.archives.cz/da/Zoomify.action?entityRef=\%28\%5En\%29\%28\%28\%28localArchiv\%2C\%5En\%2Chot\\_\%29\%28unidata\%29\%29\%28414789\%29\%29&scanIndex=0](https://digi.archives.cz/da/Zoomify.action?entityRef=\%28\%5En\%29\%28\%28\%28localArchiv\%2C\%5En\%2Chot_\%29\%28unidata\%29\%29\%28414789\%29\%29&scanIndex=0)

- [url pro procházení stránek scanů archiválie](#)

– pro přístup potřeba platná cookie

```
https://digi.archives.cz/da/PaginatorMedia.action
?_sourcePage=PxJP4Y-ARPeWynGwwTghRSxZwXUWba_Kailsppj
-Ipuxs9UsooqQCR-IjhdKCoa4IyvrlI
Bv9Y1yGNpxal2xD64DaagDoUcfqww4DZ3iIhY\%3D
&row=10
```

- [permalink scanu archiválie \(např. prvního scanu N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková\)](#)

```
https://digi.archives.cz/da/permalink
?xid=be89b34c-f13c-102f-8255-0050568c0263
&scan=f0c2acd4516c42e0872a0d1122e81f2c
```

- [zoomify složka \(např. prvního scanu N, Z, O • inv. č. 3269d • sig. JV I 5 • Bernartice, Buková\)](#)

```
https://images.archives.cz/mrimage/matriky/zoomify
/cz/archives/CZ-217000010/NAD-165/dao/images
/0416/cad1429d-f000-4936-99e3-4c7d6de383a0.jpg
```

## Státní oblastní archiv v Praze 2.5

- [vyhledávač matričních záznamů](#)

– poslouží pro získání platné cookie

```
https://ebadatelna.soapraha.cz/pages/SearchMatrikaPage
```

- [první stránka matričních záznamů](#)

– potřeba nastavit platnou cookie

```
https://ebadatelna.soapraha.cz/wicket/page
?0-1.IBehaviorListener.0-searchForm-searchButton
```

- POST

```
idd_hf_0=
searchType=1
keywordBox%3AsearchTokenIn=radio7
keywordBox%3AsearchToken=lokalita
dateYearFrom=1000
dateYearTo=2100
searchButton=1
```

- pro následující stránky matričních záznamů

– potřeba nastavit platnou cookie

```
https://ebadatelna.soapraha.cz/wicket/page
?0-1.IBehaviorListener
.1-tableContainer-pagination-results-nextPageUrl
```

- url matričního záznamu (např. Arnoštovice 01)

```
https://ebadatelna.soapraha.cz/pages/MatrikaPage/matrikaId/3726
```

- url prvního scanu maticního záznamu (např. 1. scan Arnoštovice 01)

– získání dalších scanů změnou id strany v url

```
https://ebadatelna.soapraha.cz/pages/MatrikaPage/matrikaId/3726
```

## Státní oblastní archiv v Plzni 2.6

- matriční rozcestník

– - vede na rozcestník dle územního rozsahu.

```
https://www.portafontium.eu/contents/register/
```

- stránka územního rozsahu (např. Aš - evangelická)

– pro přechod na další stranu, přidání proměnné page na konec url, počítáno od 0

```
https://www.portafontium.eu/contents
/register/soap-pn/cirkev-evangelicka/as-evangelicka
```

- archiválie (např. Aš-evangelická 01)

```
https://www.portafontium.eu/register/soap-pn/as-evangelicka-01
```

- url scanu archiválie (např. Aš-evangelická 01)

– všechny následující url lze získat rovnou ze stránky archiválie.

– ze scriptu vloženého na stránce, získat url xml IIPImage a kořenových složek dlaždič všech scanů archiválie 19 script na stránce

```
https://www.portafontium.eu
/iipimage/30060059/as-evangelicka-01_0010-x
```

- xml IIPImage scanu archiválie (např. první scan archiválie Aš-evangelická 01)

[https://www.portafontium.eu/fcgi-bin/iipsrv.fcgi  
?DeepZoom=register/soap-pn/as-evangelicka/as-evangelicka-31  
/as-evangelicka-31\\_0010-x.jp2.dzi](https://www.portafontium.eu/fcgi-bin/iipsrv.fcgi?DeepZoom=register/soap-pn/as-evangelicka/as-evangelicka-31/as-evangelicka-31_0010-x.jp2.dzi)

- kořenová složka dlaždic scanu archiválie (např. první scan archiválie Aš-evangelická 01)

[https://www.portafontium.eu/fcgi-bin/iipsrv.fcgi  
?DeepZoom=register/soap-pn/as-evangelicka/as-evangelicka-31  
/as-evangelicka-31\\_0010-x.jp2\\_files/](https://www.portafontium.eu/fcgi-bin/iipsrv.fcgi?DeepZoom=register/soap-pn/as-evangelicka/as-evangelicka-31/as-evangelicka-31_0010-x.jp2_files/)

## Moravský zemský archiv v Brně 2.7

- seznam matričních záznamů rozdělen podle původce (např. československá církev evangelická)

- původce lze měnit změnou hodnoty proměnné `typ_puvodce_id`
- přechod na další stranu přidáním, změnou hodnoty proměnné `page`
- seznam původců s příslušnými hodnotami proměnné:

- \* československá církev evangelická `typ_puvodce_id=1`
- \* československá církev `typ_puvodce_id=2`
- \* civilní matriky `typ_puvodce_id=3`
- \* Jednota baptistů `typ_puvodce_id=4`
- \* Jednota bratrská `typ_puvodce_id=5`
- \* Milosrdní bratři `typ_puvodce_id=6`
- \* německá evangelická církev `typ_puvodce_id=7`
- \* německé vojenské matriky `typ_puvodce_id=8`
- \* německý stavovský úředník `typ_puvodce_id=9`
- \* pravoslavná církev `typ_puvodce_id=10`
- \* řeckokatolická církev `typ_puvodce_id=11`
- \* římskokatolická církev `typ_puvodce_id=12`
- \* starokatolická církev `typ_puvodce_id=13`

[https://www.mza.cz/actapublica/matrika/hledani  
?typ=puvodce  
&typ\\_puvodce\\_id=1](https://www.mza.cz/actapublica/matrika/hledani?typ=puvodce&typ_puvodce_id=1)

- archiválie, odkazuje přímo na první scan (948 Olešnice na Moravě (helvetské vydání))

[https://www.mza.cz/actapublica/matrika/detail/3734  
?image=216000010-000253-003368-000000-000948  
-000000-00-B01996-00010.jp2](https://www.mza.cz/actapublica/matrika/detail/3734?image=216000010-000253-003368-000000-000948-000000-00-B01996-00010.jp2)

- xml IIPImage scanu archiválie (např. první scan archiválie 948 Olešnice na Moravě (helvetské vydání))

– pro přístup potřeba nastavit v HTTP hlavičce dotazu položku:

\* Referer: <https://www.mza.cz/>

```
https://www.mza.cz/actapublica/iipsrv/iipsrv.fcgi
?Deepzoom=/srv/LUN013/LUN013/216000010
/216000010-000253-003368-000000-000948-000000-00-B01996
/216000010-000253-003368-000000-000948-000000-00-B01996-00010.jp2.dzi
```

- [kořenová složka dlaždic scanu archivialie \(např. první scan archivialie 948 Olešnice na Moravě \(helvetské vydání\)\)](#)

– pro přístup potřeba nastavit v HTTP hlavičce dotazu položku:

\* Referer: <https://www.mza.cz/>

```
https://www.mza.cz/actapublica/iipsrv/iipsrv.fcgi
?Deepzoom=/srv/LUN013/LUN013/216000010
/216000010-000253-003368-000000
-000948-000000-00-B01996
/216000010-000253-003368-000000-000948
-000000-00-B01996-00010.jp2_files/
```

## Archiv města Ostravy 2.8

- [seznam matričních záznamů](#)

```
https://badatelna.ostrava.cz/vademecum/searchlink?fcDb=10058
```

- [permalink matričního záznamu \(např. Archiv obce Bartovice , inv. č. 20c\)](#)

```
https://badatelna.ostrava.cz/vademecum/permalink
?xid=CD9DB9B3A15311E9B75C000C29D8B843
```

- [odkaz na první scan archivialie](#)

```
https://badatelna.ostrava.cz/vademecum/permalink
?xid=CD9DB9B3A15311E9B75C000C29D8B843
&scan=7b6e1873d9fb4bf089aac6dfc2a0182a
```

- [kořenový adresář Zoomify složky](#)

```
https://badatelna.ostrava.cz/mrimage/vademecum/zoomify/cz
/archives/amo/inventare/dao/images/0001/56_20c/00001.jpg
```

## Kapitola 4

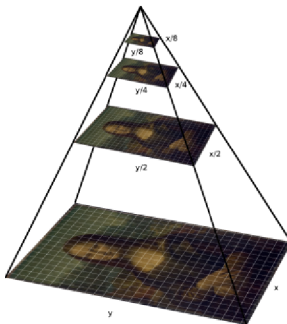
# Zobrazování obrázků ve vysokém rozlišení na webu

V kapitole se zaměřím na seznámení formáty, protokoly a aplikacemi, které používají archivy pro zobrazování obrázků s vysokým rozlišením. Obrázky ve vysokém rozlišení zpravidla bývají paměťově náročné, často se jedná o soubory s velikostí desítek MB. Načítání takových obrázků na web je tedy velmi nepraktické. Formáty obrázků a protokoly definující jak těmito soubory pracovat, tento problém řeší rozdělením obrázku do několika vrstev o různých rozlišeních. Takové obrázkové formáty jsou označovány jako *multi-resolution image format*, neboli obrazový formát s více rozlišeními. Klient si obrázek tedy nestahuje ze serveru celý, ale stahuje pouze požadované části obrázku. Takový formát obrázku, musí mít definovaný způsob jakým se vytvoří z originálního obrázku, klientskou aplikaci, která části obrázku poskládá do správného tvaru a protokol, přes který ze serveru získává požadované části.

### 4.1 Obrazové formáty s více rozlišeními

#### 4.1.1 Tiled Multi-Resolution TIFF

Je rozšířením obrazového formátu TIFF, který se používá pro ukládání rastrové grafiky. V tomto rozšíření jsou jednotlivé rozlišení uložena v oddělených vrstvách. Tento způsob bývá prezentován jako maticová pyramida.



Obrázek 4.1: Znázornění maticové pyramidy rozlišení[8]

### 4.1.2 JPEG2000

Je formát pro kompresi obrázků založený na vlnkové transformaci. Při kompresi obrazu dochází k rozkladu obrazu do různých rozlišení, vzniká tedy maticová pyramida, kterou lze využít i mimo kompresi.

## 4.2 Protokoly a aplikace

### 4.2.1 IIPImage

IIPImage se skládá ze serverové aplikace, protokolu IIP a webového klientu IIPMooVier. Vydáný jako OpenSource pod GNU General Public Licension [7].

#### Server

Serverová aplikace pro streamování obrázků ve vysokém rozlišení, podporuje obrazové formáty s více rozlišeními TIFF i JPEG2000. Program je napsán v jazyce C++. Pro komunikaci s lokálním webovým serverem jako je Apache nebo Nginx je potřeba provést správnou konfiguraci. Je multiplatformní, u některých linuxových distribucí je součástí oficiální balíčků (např. Ubuntu).

Server umí zpracovávat požadavky různých protokolů. Jedná se především o protokol IIP, který je výchozím protokolem. Dalšími protokoly jsou IIF, Zoomify, DeepZoom.

#### Protokol IIP

Internet Imaging Protocol. Rozhraní protokolu je založené na REST, tedy pomocí HTTP protokolu a jeho základních metod (GET, POST, PUT, DELETE) lze přistupovat k jednotlivým dlaždicím, metadatům nebo příkazům nad obrázkem.

Příklad základních příkazů, všechny následující příkazy jsou předávány HTTP metódou GET, argumenty je většinou zapotřebí kombinovat.

- `FIF=path` plná cesta k obrázku na serveru
- `JTL=r,n` získání dlaždice `n`-té při rozlišení `r` ve formátu JPEG
- `obj=IIP,1.0&obj=Max-size&obj=Tile-size&obj=Resolution-number` pro získání metadat o brázku

#### Webový klient IIPMooViewer

Webový klient založený na technologiích HTML5, CSS3 a JavaScript. Umí kromě komunikace IIP protokolem i další protokoly IIF, Zoomify a další.

### 4.2.2 Zoomify

Je společnost poskytující webového klienta, protokol a program pro tvorbu dlaždic. Produkty jsou vydány v několika verzích, zdarma a placené. Obrázky musí být před umístěním na webový server zpracovány do dlaždic. Dlaždice jsou umístěné v adresářové struktuře přímo na webovém serveru. Výhodou tohoto přístupu, je vypuštění speciální aplikace přímo na serveru. Klient tedy přistupuje přímo k jednotlivým dlaždicím. Nevýhodou, takového uložení dlaždic může být větší náročnost na paměť.

## Protokol Zoomify

Jelikož jsou dlaždice přímo uloženy na webovém serveru, protokol je definován adresářovou strukturou. Je tedy potřeba znát kořenovou složku obrázku `img_root`. Kořenová složka je dělena do skupin dlaždic `TileGroup $x$` , kde  $x$  značí skupinu.

- `img_root/ImageProperties.xml` získání metadat o obrázku
- `img_root/TileGroup0/r-x-y.jpg` získání dlaždice ze skupiny 0 ve vrstvě  $r$  na pozici  $x,y$

### 4.2.3 OpenSeadragon

Je OpenSource webový klient založený na technologiích HTML5, CSS3 a JavaScript. Není vázán na konkrétní protokol. Mezi protokoly které umí se řadí následující *IIP*, *IIF*, *Zoomify*, *DeepZoom*.

### Další protokoly

Existují i další protokoly, které zde uvedu bez podrobnějšího popisu. Protože nejsou využívány archivy u nás.

- DZI – DeepZoom
- IIF – International Image Interoperability Framework



# Kapitola 5

## Návrh webu

V této části se věnuji návrhu vlastního webu pro zobrazování archiválií. V první řadě se bude jednat o návrh uživatelského rozhraní celého webu a především části, která bude sloužit k zobrazování archiválií. Dále se v této kapitole zaměřím na návrhu databáze, databáze bude sloužit k uchování informací o jednotlivých archivech, kategoriích archivů, záznamů a scanů. Další důležitá část návrhu se týká scriptů pro automatické získávání informací a jejich propojením se zbytkem webu.

### 5.1 Návrh uživatelského rozhraní

Při návrhu uživatelského rozhraní jsem kladl důraz především na jednoduchost a přehlednost výsledného webu.

#### Domovská stránka

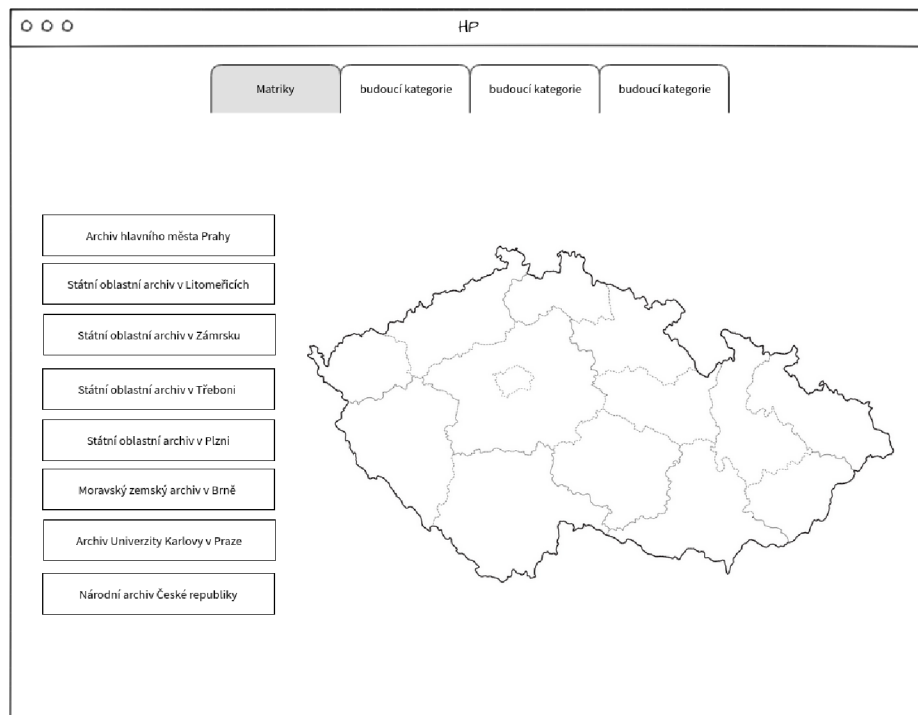
Grafické rozvržení je vidět na wireframu 5.1. Na stránce je seznam archivů, dostupných kategorií. V návrhu je vidět i mapa ČR, která slouží k zobrazení územního rozsahu jednotlivých archivů. Mapa slouží zároveň jako odkaz do přehledu vybrané kategorie (horní navigační lišta) a archivu (postranní navigační panel).

#### Přehled kategorie archivu

Na obrázku 5.2 je vidět grafický návrh sekce, sloužící pro výpis jednotlivých záznamů ve vybrané kategorii archivu. V horní navigační liště lze změnit prohlíženou kategorii. Důležitým prvkem této sekce je vyhledávání v kategorii. Vyhledávací filtry jsou odvozeny od atributů sekce. Na stránce jsou zobrazeny nalezené záznamy. Po rozkliknutí záznamu se na stránku propíše i informace o konkrétním záznamu. Ze záznamu vede odkaz na zobrazení jednotlivých scanů archiválie.

#### Zobrazení archiválie

Návrh je vidět na obrázku 5.3. Jedná se část webu, na které uživatel stráví většinu času. Z toho důvodu musí být práce s galerií co nejpřívětivější. Hlavní část stránky zabírá sekce pro zobrazení obrázku. Důležité je pohodlné ovládání této části. To by mělo zahrnovat posouvání zobrazované oblasti myší i klávesami, přiblížení nebo oddálení při scrollování nad touto oblastí i klávesami. Po stranách jsem umístil tlačítka pro přechod na další, předchozí obrázek. Umístění těchto tlačítek jsem zvolil především z důvodu stejného umístění



Obrázek 5.1: Návrh uživatelského rozhraní webu, domovská stránka

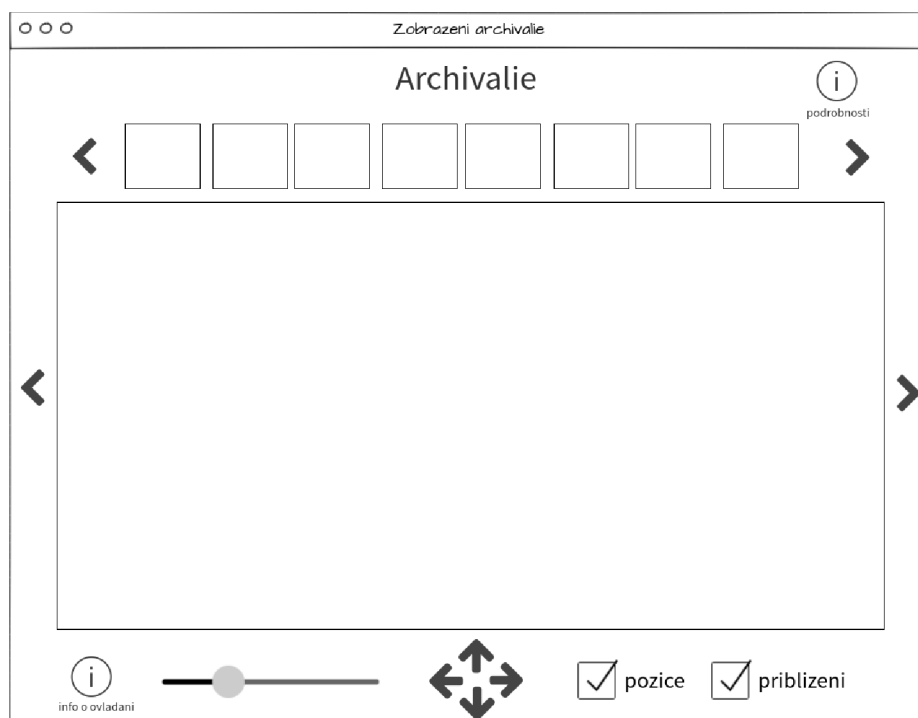


Obrázek 5.2: Návrh uživatelského rozhraní webu, stránka archivu

u většiny obrázkových galerií se kterými se uživatelé setkávají (např. Facebook). Přejít na další obrázek se provede bez opětovného načtení stránky.

Nad hlavní částí se nachází lišta zobrazující náhledy dalších obrázků. Přejít na další stránku náhledů je možné pomocí tlačítek postranách lišty. Zobrazení dalších náhledů se provede bez opětovného načtení celé stránky, stejně jako při přechodu na další zobrazované obrázky. Náhledy slouží i ke změně zobrazovaného obrázku. Aktivní obrázek bude v náhledech zvýrazněn. Umístění lišty náhledů nad zobrazovaný obrázek jsem zvolil, především kvůli prohlížení na mobilu.

Pod hlavní částí je ovládací lišta. Tato lišta obsahuje posuvník, pomocí kterého lze měnit přiblížení. Dalším prvkem jsou tlačítka sloužící pro posun zobrazované části obrázku. Následují checkboxy pro nastavení zda se bude držet pozice a přiblížení při přechodu na další obrázek. Důležitým prvkem je tlačítko pro zobrazení pop up okna s popisem ovládaní.

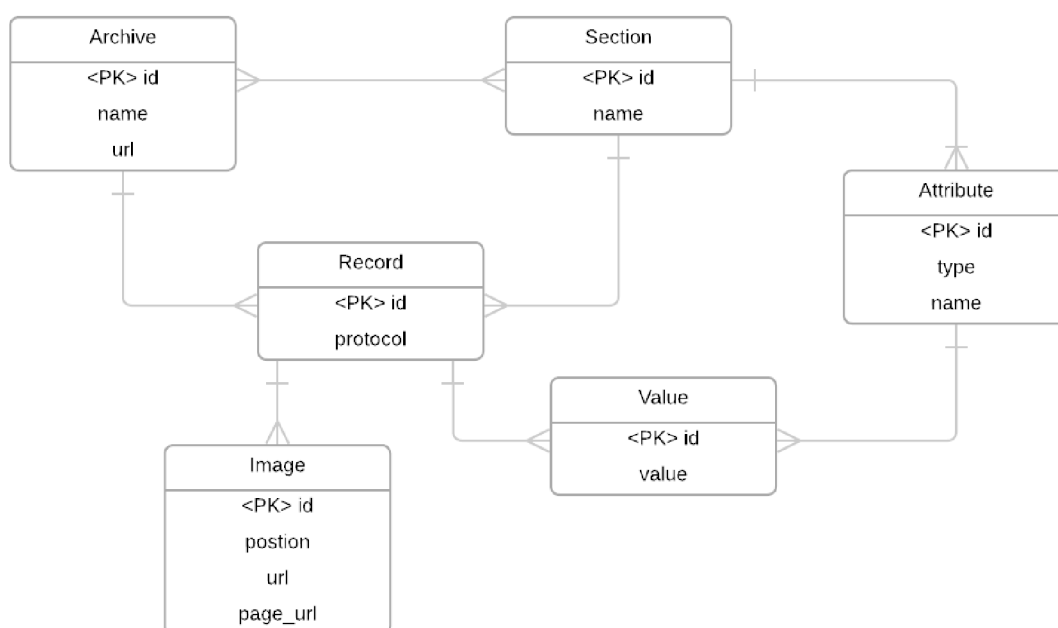


Obrázek 5.3: Návrh uživatelského rozhraní webu, stránka zobrazení archiválie

## 5.2 Návrh databáze

Návrh databáze ve formě ERD je k nahlédnutí na obrázku 5.4. Návrh databáze obsahuje entity *Archive*, *Section*, *Attribute*, *Value*, *Record* a *Scan*.

Entita *Archive* reprezentuje archiv, uchovává data jako je jméno, odkaz na web archivu. Případně i informace o území, které archiv pokrývá. Při návrhu entity pro uchování informací o jednotlivých kategoriích archivu, *Section*, jsem se musel zamyslet nad způsobem jak v budoucnu přidávat sekce, neboli přibudou sekce u kterých nevím počet ani datový typ jejich vlastností. Tento zádrhel jsem vyřešil zavedením entity *Attribute*. Tato entita obsahuje pouze dva sloupce, jméno a datový typ. Sloupec datový typ bude omezen předem definovanou množinou konstant. Konkrétní záznam je v návrhu zastoupen entitou *Record*, ten je vždy provázán s jediným archivem a jedinou sekcí. Každý záznam je provázán s entitami typu *Value*, které uchovávají hodnoty záznamů. Entita *Value* je vždy provázána entitou *Attribute*.



Obrázek 5.4: ER diagram návrh schéma databáze

## 5.3 Automatizované získávání obsahu

Srdcem celého webu jsou data z archivů, bez nich web nemá žádný reálný přínos. Proto je potřeba navrhnout skripty, které nejen že web inicializují daty, ale i budou udržovat informace aktuální. Bohužel tyto skripty se budou muset implementovat pro každou kategorii i každý archiv zvlášť. Jejich činnost je definována podobným algoritmem. Skripty rozdělíme na části podle typu dat, který získávají na následující kategorie.

- skripty pro extrahování a uložení seznamu všech záznamů kategorie
- skripty pro extrahování a uložení informací jednotlivých záznamů

- skripty pro extrahování a uložení všech scanů záznamu

Každý z těchto skriptů potřebuju vstuní stránku, selector vybíraného obsahu, odkaz pro přechod na další stránku, nebo selector jak tento odkaz získat. Obecný algoritmus pro získání scanů matričních záznamů je uveden 1.

---

**Algorithm 1** Obecný algoritmus pro získání matričních záznamů archivu

---

```

1: procedure ARCHIVEDATAMINE(urlMatrRozcestnik)
2:   urlMatrik = []
3:   while urlMatrRozcestnik ≠ null do
4:     pageHTML = obsah z url urlMatrRozcestnik
5:     urlsFromPage = vyber url matrik z pageHTML
6:     pridej urlsFromPage do urlMatrik
7:     urlMatrRozcestnik = ziskej url nasledujici stranky urlMatrRozcestnik
8:   for all url ∈ urlMatrik do
9:     scanUrls = []
10:    tileFoldersUrls = []
11:    pageHTML = obsah z url matriky url
12:    galeryUrl = vyber url prohlizece scanu z pageHTML ▶ 1.scan
13:    while galeryUrl ≠ null do
14:      scanHTML = obsah z url galeryUrl
15:      scanPermalink = vyber permalink scanu z scanHTML
16:      pridej scanPermalink do scanUrls
17:      tileUrl = vyber url slozky dlazdic obrazku z scanHTML
18:      pridej tileUrl do tileFoldersUrls
19:      galeryUrl = vyber url nasledujiciho scanu z scanHTML
20:    vloz do databaze url, scanUrls, tileFoldersUrls

```

---

## 5.4 Návrh prohlížeče archiválií

Při návrhu prohlížeče archiválií jsem musel vzít v potaz technologie, které jsou využívány archivy. A zajistit jednotné rozhraní pro práci s těmito různými klienty a protokoly. Pro shrnutí uvádím přehled jaké technologie jsou používány jednotlivými archivy v tabulce ???. Z průzkumu tedy vyplývá, že můj prohlížeč musí být schopen pracovat s protokoly Zoomify a OpenSeaDragon. Dále je potřeba zobrazování obrázků, které nejsou ve multi-resolution formátu. Pro práci s archivy využívající Zoomify a OpenSeaDragon, bych mohl použít některou z knihoven podporující různé formáty např. IIP nebo OpenSeaDragon, obě tyto služby jsou OpenSource a rozumí požadovaným protokolům.

Dále je potřeba rozhodnout jak pracovat s archivy, které obrázky neposkytují v žádném multi-resolution formátu. Jedním řešením by mohlo být pro takové obrázky udělat vlastní prohlížeč, který bude pracovat s obrázky v plném rozlišení. Jako další řešení by se dal využít serverová aplikace IIP, při požadavku na takový obrázek, by nejdříve musel server obrázek stáhnout k sobě do mezipaměti, konvertovat ho na některý z multi-resolution formátů, otázkou však zůstává jak časově náročné tato konverze bude, nabízí se i otázka jak poznat, že obrázek lze z mezipaměti odstranit. V tomto řešení by klient mohl být stejný jako u ostatních archivů.

## Kapitola 6

# Implementace

V kapitole popisují jednotlivé kroky a rozhodnutí, kterými jsem procházel při implementaci vlastního řešení. V kapitole se nejprve věnuji přehledu použitých technologií a zhodnocení proč jsem se rozhodl právě pro tyto technologie. Dále popisují tvorbu klientské části *front-end*, především tedy stránky pro zobrazování digitalizovaných záznamů archiválií. Následně se zaměřím na serverovou část *back-end*. V té se jedná především o samotnou strukturu celého webu, jednotlivé přístupové body a jejich propojení s *front-endem*. Nedílnou součástí je správa a komunikace s databází. V neposlední řadě pak také skripty pro získávání dat z archivů. Při implementaci jsem narazil i na řadu, problémů které jsem při návrhu nevzal v potaz. Proto se zaměřím na porovnání důvodu proč taková situace nastala, jak jsem jí odstranil a odchylku implementace od návrhu. Pokusím se uvést jak se takovým situacím při budoucích návrzích vyvarovat.

### 6.1 Použité technologie

Před samotnou implementací jsem se musel rozhodnout jaké technologie použiji. Co se týče klientské části - *front-endu*, není ani nad čím uvažovat, jedná se o technologie HTML5, CSS3 a JavaScript. Otázka však nastává, zda se vyplatí využít některý z JavaScriptových frameworků pro tvorbu dynamických uživatelských rozhraní. Po zvážení jsem došel k závěru, že minimálně při tvorbě stránky pro zobrazování archiválií by mohl ulehčit implementaci. Vybrat vhodný JavaScriptový framework v dnešní době není nic snadného, existuje je jich spousta např. Angular.js, React.js, Svelte, Vue.js. Kvůli lehkým předchozím zkušenost s pouze jediným z vyjmenovaných frameworků jsem se rozhodl pro React.js. Při implementaci jsem ovšem narazil na problémy, které podrobněji zmíním níže a přesvědčili mě ke změně pro Svelte.

Základní technologie pro tvorbu klientské strany webu mám vybranou. Před přesunutím se na výběr *back-endových* technologií je potřeba ještě učinit finální rozhodnutí ohledně zobrazování archiválií. Především jakou knihovnu použít. Po důkladném zhodnocení požadavků od jednotlivých archivů jsem se rozhodl využít webového klienta IIPMooViewer.

Při volbě technologií pro *back-end* jsem se musel rozhodnout, zda použiji některý z webových serverů jako je Apache, Nginx nebo jestli se rozhodnu napsat si vlastní server napsaný např. v node.js nebo Pythonu. Pro typ aplikace webové aplikace, jako tato mi přišlo rozumnější vybrat některý z vysoce odzkoušených serverů než tvořit vlastní. Jak se říká nevyráběj znovu kolo. Vybral jsem si tedy jako webový server Apache. Jako programovací jazyk jsem zvolil PHP 8.0 především, kvůli jeho praktičnosti při tvorbě webových

aplikacích. Back-endový stack jsem doplnil o framework Symfony pro PHP, hlavním důvodem pro použití frameworku bylo snadná implementace routování, bez nutnosti tvořit pro každý přístupový bod soubor v adresářové struktuře na serveru, ale i šablonovací systém Twig. Pro databázi jsem zvolil MySQL, hlavním důvodem byly mé zkušenosti s touto databází.

## Svelte

Svelte je nástroj pro tvorbu uživatelských rozhraní. Hlavním rozdílem Svelte od jeho konkurentů jako React.js, Vue.js a Angular.js je že se nejedná o JavaScriptový framework. Zdrojový kód ve Svelte se přímo nekládá do webové stránky, ale nejdříve se zkompiluje do JavaScriptu, až zkompilovaný kód se vkládá do webové stránky. Jedná se o jakousi nadmnožinu jazyka JavaScriptu doplněnou o syntaxi definující reaktivní chování. Frameworky jako React nebo Vue používají k zpracování reactivity aplikace virtuální DOM. Reaktivitu si lze představit jako provázanost elementů, tedy změna některého prvku je automaticky propagována a nemusí to řešit programátor například zaváděním návrhových vzorů jako je Observer. Svelte nevyužívá virtuální DOM pro uchování stavu UI, ale tyto závislosti jsou detekovány během překladač a nahrazeny funkcemi, které přímo upravují DOM při změně stavu. [15]

## PHP

PHP je scriptovací jazyk používaný především k tvorbě webových aplikací. Vykonává se na straně serveru. Odhaduje se, že i v současnosti běží zhruba 70% webových aplikací právě na PHP [19]. Výhoda PHP při tvorbě webových aplikací je, že HTML kód lze vkládat přímo do zdrojového kódu PHP, od toho přístupu se ovšem v posledních letech upouští, protože tak dochází ke snížení čitelnosti a přehlednosti programu. Tento přístup bývá nahrazen návrhovým vzorem *MVC – Model View Controller*. Existují frameworky jako Symfony, Nette, Laravel a další, které poskytují základní strukturu MVC.

## Symfony

Symfony je Open Source PHP framework, složený z množiny znovupoužitelných komponent. Pomáhá v rychlejší vývoji a tvorby strukturovaného zdrojového kódu. Tím je zajištěna lepší dlouhodobá udržitelnost projektu a rozšiřovatelnost. Uvedu některé důležité komponenty, kvůli kterým jsem se rozhodl pro využití toho frameworku.

- Routing, AbstractController [17]
- Doctrine ORM [16]
- HttpBrowser, HttpClient, DomCrawler [18]

Podrobné seznámení s vyjmenovanými komponenty uvedu níže v textu. Při popisu implementace back-endu.

## 6.2 Serverová část

### Adresářová struktura aplikace

Symfony definuje adresářovou strukturu následovně. Veškeré autorské zdrojové kódy serverové části aplikace se nacházejí v adresářích `src/` a `template/`. Přístupovým bodem celé

aplikace je `public/index.php`, proto se při konfiguraci serveru musí adresář `public/` definovat jako `DocumentRoot`. V adresáři `public/` se dále mohou definovat CSS a JavaScript soubory, které si klient stahuje při načítání stránky. Dále jsou v kořeni adresáře a soubory potřebné pro správný chod Symfony. Dalším důležitým adresářem je `vendor/`, tento adresář je generován programem `composer`, což je správce balíčků pro PHP, v tomto adresáři jsou zdrojové kódy veškerých potřebných balíčků třetích stran. Seznam použitých balíčků je v souboru `composer.json`.

## Routování a kontrolery

Kontrolery jsou definovány v adresáři `Controller/`, routy jsou definovány u příslušných metod kontroleru. Routy se v Symfony zapisovaly pomocí anotací viz. 6.2, od verze PHP 8.0 je můžeme zapisovat zapomocí nativních atributů 6.2.

```
class RecordController
{
    /**
     * @Route("/record/{id}", name="record")
     */
    public function get($id) { /* ... */ }
}

class RecordController
{
    #[Route(path: '/record/{id}', name: 'record')]
    public function get(int $id) { /* ... */ }
}
```

Kontrolery a jím příslušné routy jsou definovány následovně. Reflektují entity, jejichž data uživatel chce zobrazit.

- `RecordController`
  - `/record/{id}` - detail záznamu
  - `/record/{id}/gallery` - prohlížeč scanů záznamu
- `ArchiveController`
  - `/archive/{id}` - přehled archivu, jeho kategorií
  - `/archive/{archiveId}/section/{sectionId}` - kategorie archivu, seznam záznamů
- `BrigeController`
  - `/bridge/{data}` - přistupuje na url v proměnné data a vrací odpověď
- `HomeController`
  - `/` - home page



Po přijetí požadavku, dochází k jeho zpracování. Téměř vždy jsou k tomu zapotřebí data z databáze, v takovém případě požádá kontroler `EntityManager` – instanci ORM, o získání konkrétní entity z databáze, nebo v případě složitějších dotazů o získání příslušného repozitáře, ve kterém jsou definovány konkrétní dotazy pro získání celých kolekcí entit. Kontroler následně předává zpracovaná data Twig šabloně, která se vrací jako odpověď.

## Entity a repozitáře

Entity jsou třídy definující jednotlivé tabulky a relace databáze. Po definici nové entity, nebo upravě některého atributu, je potřeba změny propagovat do databáze. V Symfony se tato akce označuje jako migrace. Migrace se spouští dvojicí příkazů.

```
$ php bin/console doctrine:migrations:diff
$ php bin/console doctrine:migrations:migrate
```

Tyto příkazy se tedy pouze volají při změně schéma databáze.

Každá třída definující entitu musí být provázána s repozitářem. Repozitář je třída rozšiřující, třídu `Doctrine\ORM\EntityRepository`. Repozitář definuje dotazy pro získání více než jednoho řádku příslušné tabulky. V repozitáři se definují vlastní dotazy. Při změně dat v entitě je potřeba tyto změny uložit do databáze. O to se stará `EntityManager`, entita se nejprve musí uložit do mezipaměti zavoláním metody `persist`, nad instancí `EntityManager`. Tato funkce přijímá jediný argument, instanci entity. Vyprazdnění mezipaměti a vložení dat do databáze se uskuteční voláním metody `flush`, nad instancí `EntityManager`. Entity jsou definovány v adresáři `Entity/`, repozitáře v adresáři `Repository/`.

Seznam entit, atributy entit jsou uvedé v návrhu databáze. Vazby mezi entitami se definují přímo v entitě. Ukázka N:1 vazby mezi entitami `Value` a `Record` níže [6.2](#).

```
#[Entity(repositoryClass: RecordRepository::class)]
class Record{
    ...

    #[ORM\OneToMany(
        targetEntity: Value::class,
        mappedBy: 'record',
        cascade: ['persist', 'remove']
    )]
    private Collection $values;
}

#[Entity(repositoryClass: ValueRepository::class)]
class Value{
    ...

    #[ORM\ManyToOne(
        targetEntity: Record::class,
        inversedBy: 'values'
    )]
    private Record $record;
}
```

## Příkazy

V adresáři `Commands/` jsou definovány třídy rozšiřující `Symfony\Component\Console\Command`, které se dají spouštět pouze přes terminál. PHP skripty spuštěné v terminálu, nejsou omezené časovým limitem jako skripty spuštěné na serveru. Aby v těchto skriptech bylo možné využít Doctrine ORM, je potřeba definovat je pomocí těchto tříd. V aplikaci slouží ke spuštění skriptů pro získávání dat z archivů, nebo pro přidání nových archivů nebo sekcí. Ukázka definice příkazu a jeho následné volání [6.2](#). Je třeba implementovat metodu `execute`. Takto definované příkazy je poté možné volat např. pomocí `Cron` a spouštět je periodicky.

```
#[AsCommand(
    name: 'app:test',
    description: 'test command',
    hidden: false,

)]
class CommandTest extends Command
{
    protected function execute
    (
        InputInterface $input,
        OutputInterface $output
    ): int
    {
        //tělo příkazu
    }
}
```

Volání příkazu je následovné.

```
$ php bin/console app:test
```

## Skripty pro automatizované získávání dat

V adresáři `Scraper/` je připravena abstraktní třída `AbstractScraper`, třídy rošiřující tuto třídu musí implementovat metodu `process`, ve které se definuje konkrétní chování. Třída vytváří instanci `Symfony\Component\BrowserKit\HttpBrowser`, pomocí které simuluje chování webového prohlížeče. Pomocí této třídy je možné odesílat HTTP požadavky, klikat na odkazy nebo získat obsah stránky. Tyto skripty se musí registrovat do některého z `Commandu` a spouštět pouze přes něj. Pro extrakci dat jednotlivých kategorií archivů, bude potřeba implementovat vlastní scraper, specializovaný na konkrétní archiv a kategorii. Bohužel z nedostatku času jsem implementoval pouze script pro získání všech scanů jednoho matričního záznamu Zemského archivu v Opavě [2.3](#).

Při testování jsem narazil, na dobu která je potřebná k extrakci url jednotlivých scanů jednoho záznamu. Tato operace v průměru byla v jedntotkách sekund. Lze si tedy představit, že zpracování velkého množství záznamů potrvá věky. Toto je místo pro optimalizaci.

## 6.3 Klientská část

Nejdůležitější komponentou na klientské části je prohlížeč digitalizovaných scanů archiválie. Uživatelské rozhraní je této části je kombinace Svelte a knihovny `IIPMooVier` [4.2.1](#). Uživa-

telské rozhraní jsem původně začal psát v React.js, ale kvůli potížím propojení knihovny IIPMooVier s Reactem, jsem vyzkoušel Svelte a problém byl hned vyřešen. Výhoda Svelte je, že zkompileovaný script je soubor funkcí. Tudíž stačilo do HTML kódu stránky prvně vložit scripty IIPMooVier.

Knihovnu IIPMoovier jsem zvolil kvůli počtu protokolů, které podporuje. A snadného doplnění dalšího protokolu, bez nutnosti zásahu do knihovny. Zdrojové kódy knihovny jsem musel rozšířit o některé funkce. Jedná se především o funkci `getMetaDataURLWrapper`, tuto funkci jsem musel zavést společně s atributem `bridge`, v tomto atributu je definována url `bridge` přes, který knihovna přistupuje k souboru s metadaty o obrázku. `Bridge` jsem musel zavést jelikož knihovna k souboru přistupuje pomocí AJAX, ovšem soubor se nachází na jiné doméně a cizí server, nemá nakonfigurované HTTP Cors hlavičky. Prohlázeč takové dotazy nepovoluje. Použitím `bridge`, který se dotáže serveru a odpověď předá klientovy, se problém řeší. Toto řešení ovšem není ideální a otevírá potencionální bezpečnostní slabiny.

Další částí knihovny, kterou jsem musel upravit byl protokol Zoomify, protokol byl nejspíše od doby vydání knihovny aktualizován.

Knihovna IIPMooVier je obalena Svelte komponentami, které jsou definovány v adresáři `public/svelte/src/` a které řídí reagují na uživatelské akce, které předávají knihovně. Na nejvyšší úrovni je to komponenta `Viewer`, tato komponenta přijímá dva argumenty, prvním je pole adres obrázků, druhý je objekt IIPMooViewer. Oba tyto argumenty jsou vygenerovány na serveru. Neodchází tudíž k žádným dalším inicializačním načítáním. Komponenta `Viewer`, uchovává ukazatel na uaktuální obrázek. Změnou hodnoty ukazatele, dochází i ke změně zobrazovaného obrázku. Vnitřek komponenty je tvořen třemi částmi, hlavičkou která je tvořena komponentou `Previews` obsahující miniatury dalších snímků. Miniaturami je možné se pohybovat, bez načítání stránky. Při přechodu na další stranu, miniatur se načítají další obrázky. Pokud se uživatel pohybuje v rámci jedné stránky, nedochází k opětovnému načítání. Další komponenta, tvořící hlavní část celé rozhraní je tělo, které obsahuje plátno na kam se vykresluje zobrazovaný obrázek, postranách plátna jsou tlačítka pro pohyb mezi obrázky. Přechod na jiný obrázek, automaticky posune i náhledy. Poslední částí je patička, která obsahuje tlačítka pro pohodlnější ovládání.

## Shrnutí

V kapitole jsem uvedl seznámení s použitými technologiemi a základní informace o jejich využití. Prošli jsme architekturu celé aplikace, od jejích přístupových bodů – rout, přes databázovou vrstvu až po klientskou část. Při implementaci, jsem narážel především na špatné časové rozvržení práce. To bych si vzal jako největší poučení do příštích projektů, udělat si podrobný časový plán s dostatečnými rezervami v jednotlivých blocích i na celkové doladění. Při implementaci jsem nestihl, některé prvky dodělat tak jako jsem je definoval v návrhu, jedná se především o vyhledávání záznamů v rámci sekce.

## Kapitola 7

# Testování

Testování je nedílnou součástí vývoje jakéhokoliv systému či aplikace. Neprováděl jsem žádné automatizované testy, v tomto případě je minimum přístupových bodů, takže automatizované testy, jsem nahradil testováním při vývoji.

Důkladněji jsem se ovšem věnoval uživatelskému testování. Toto testování jsem provedl na malé, ale za to velmi rozmanité skupině uživatelů. Jednalo se blízké rodinné příslušníky a spolubydlicí. Testování probíhalo ve dvou fázích. První fází jsem testoval pouze grafický návrh uživatelského rozhraní, abych odhalil případné chyby v návrhu ještě před samotnou implementací.

Testování návrhu probíhalo následujícím způsobem, testovacím uživatelům jsem předložil wireframy uživatelského rozhraní. Nedával jsem jim k nim žádné informace. Poté jsem se jim zadal jednoduché úkony a sledoval jestli v rozhraní orientují snadno a rychle. Příklady testovacích úkolů.

- Přejdi na stránku archivu  $x$
- Vyhledej záznamy v archivu  $x$  z kategorie  $y$
- Přejdi na  $n$ -tý obrázek záznamu  $x$
- Přibliž zobrazovaný obrázek
- Zobraz levý dolní roh obrázku v největším přiblížení

Při této fázi testování nebyly odhaleny žádné nedostatky v návrhu. Uživatelé ani neměli žádné připomínky. To mohlo být zapříčeno špatným způsobem testování, nebo dobrým návrhem.

V dalším kole testování jsem stejné skupině testovacích uživatelů nechal provést stejné úkoly. Bohužel toto testování probíhalo v době, kdy nebyla dokončená kompletní implementace webové aplikace. Chyběla možnost vyhledávání. Zaměřil jsem se proto především na otestování sekce pro zobrazování archiválií. V tomto testování byl odhalen, problém který jsem ve výsledné implementaci bohužel nestihl opravit. Při rychlém procházení miniatur, se v některých případech stalo, že server archivu nevrátil v požadavku žádný obrázek. Tuto situaci jsem neošetřil. Možným řešením by mohlo být přidání `EventListener`, který bude kontrolovat zda došlo k nahrání v určitém časovém intervalu, pokud ne pokusí se obrázek stáhnout znovu. Z toho testování vzešla i velmi užitečná připomínka, přidání přechodu na další snímek pomocí kláves.

## Kapitola 8

# Závěr

Obsah práce by se dal shrnout do třech základních bodů. Prvním bodem je průzkum jednotlivých archivů v České republice, během této části se ze mě stal badatel. Jako uživatel jsem si všímal věcí, které mě v bádání brzdí, většinou na straně uživatelského rozhraní. Často se jednalo o pouhé drobnosti, jejichž doplnění by pozvedlo uživatelskou zkušenost na novou úroveň. Z nejčastějších problémů jenom zmíním nemožnost návratu na předchozí stránku v katalogu záznamů, při prohlížení konkrétních scanů archiválie to zase bylo poměrně dlouhé načítání celé stránky při pohybu mezi snímky a mnoho dalšího, které jsou podrobněji rozvedeny v práci.

Druhým bodem byl návrh a implementace vlastního řešení. V této části jsem navrhoval celou architekturu aplikace, od uživatelského rozhraní, datového modelu, přístupových bodů a aktualizčních scriptů. Během návrhu datového modelu, jsem udělal chybu, když jsem se snažil definovat kategorie kompletně abstraktně, aby se v budoucnu nemusel datový vůbec měnit. Při implementaci jsem zjistil, že to nebylo nejvhodnější způsob řešení. Díky tomuto návrhu, je sice datový model pružnější, ale ztrácí na efektivnosti, což by se v budoucnu mohlo projevit například při vyhledávání. Navrhoval bych tedy tvorbu nových entit dle konkrétních potřeb. Při implementaci aktualizčních scriptů, jsem zjistil jak časově náročné je dolování dat z webu. V tomto bodě vidím příležitost do budoucna, pro zavedení paralelního zpracování. Tím by se mohl zkrátit čas, při naplnění či aktualizaci dat na použitelný čas.

Třetím a posledním bodem je testování. Z výsledku testování vyplynula pozitivní zpětná vazba od uživatelů. Bodem na kterém se testovací uživatelé shodli je sjednocení a „oživení“, vzhledu aplikace. Aplikace prý působí moc stroze.

Při celkovém zhodnocení práce, považuji všechny cíle za splněné. Možné rozšíření, nebo vylepšení do budoucna bych viděl v aktualizčních scriptech, které jsou v současném stavu spíše vhodné pro naplnění testovacími daty, než pro použití na produkci.

# Literatura

- [1] *Archivní katalog hlavního města Prahy - sbírka matrik*. [cit. 2021-29-11]. Dostupné z: <http://katalog.ahmp.cz/pragapublica/permalink?xid=7EF18906B65E11DF820F00166F1163D4>.
- [2] *Archivní katalog hlavního města Prahy - sbírka matrik*. [cit. 2021-29-11]. Dostupné z: <http://katalog.ahmp.cz/pragapublica/permalink?xid=7EF18906B65E11DF820F00166F1163D4>.
- [3] *Bavorsko-česká síť digitálních historických pramenů*. [cit. 2021-29-11]. Dostupné z: <https://www.portafontium.eu>.
- [4] *Matriční rozcestník* [online]. Česká genealogická a heraldická společnost v Praze [cit. 2021-27-11]. Dostupné z: <http://www.genealogie.cz/aktivity/digitalizace/>.
- [5] *Digitální archiv Zemského archivu v Opavě*. [cit. 2021-29-11]. Dostupné z: <https://digi.archives.cz/da/searchlink?fcDb=10041>.
- [6] *Digitální badatelna AMO*. [cit. 2021-29-11]. Dostupné z: <https://amo.ostrava.cz/digitalni-badatelna/digitalni-badatelna-a-prirucka/>.
- [7] *IIP - server*. [cit. 2021-29-11]. Dostupné z: <https://iipimage.sourceforge.io/documentation/server/>.
- [8] *IIP - images*. [cit. 2021-29-11]. Dostupné z: <https://iipimage.sourceforge.io/documentation/images/>.
- [9] *Matriky uložené v Moravském zemském archivu v Brně*. [cit. 2021-29-11]. Dostupné z: <https://www.mza.cz/actapublica/matrika/hledani>.
- [10] *Nírodní archiv archivní VadeMeCum*. [cit. 2021-29-11]. Dostupné z: <https://digi.archives.cz/da/searchlink?fcDb=10041>.
- [11] *Státní oblastní archiv v Litoměřicích - Matriční databáze*. [cit. 2021-29-11]. Dostupné z: <http://vademecum.soalitomerice.cz/vademecum/searchlink?fcAr214000010-0&fcDb=10041>.
- [12] *Státní oblastní archiv v Litoměřicích - Aplikace Archivní VadeMeCum*. [cit. 2021-29-11]. Dostupné z: <http://vademecum.soalitomerice.cz/vademecum/>.
- [13] *Státní oblastní archiv v Litoměřicích eBadatelna*. [cit. 2021-29-11]. Dostupné z: <https://ebadatelna.soapraha.cz/pages/IndexPage>.

- [14] *Státní oblastní archiv v Litoměřicích matriky*. [cit. 2021-29-11]. Dostupné z: <https://ebadatelna.soapraha.cz/pages/SearchMatrikaPage>.
- [15] *Svelte - Cybernetically enhanced web apps*. [cit. 2021-29-11]. Dostupné z: <https://svelte.dev/>.
- [16] *Symfony DOCS Databases and the Doctrine ORM*. [cit. 2021-29-11]. Dostupné z: <https://symfony.com/doc/current/doctrine.html>.
- [17] *Symfony DOCS Routing*. [cit. 2021-29-11]. Dostupné z: <https://symfony.com/doc/current/routing.html>.
- [18] *Symfony DOCS The BrowserKit Component*. [cit. 2021-29-11]. Dostupné z: [https://symfony.com/doc/current/components/browser\\_kit.html](https://symfony.com/doc/current/components/browser_kit.html).
- [19] *Usage statistics of PHP for websites*. [cit. 2021-29-11]. Dostupné z: <https://w3techs.com/technologies/details/pl-php>.