

Sem vložte zadání Vaší práce.



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE  
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ  
KATEDRA SOFTWAREVÉHO INŽENÝRSTVÍ



Bakalářská práce

## Systém pro správu elektronických verzí literárních děl

*Martin Melichar*

Vedoucí práce: ing. Karel Klouda, Ph.D

4. května 2018



---

## Poděkování

Doplňte, máte-li komu a za co děkovat. V opačném případě úplně odstraňte tento příkaz.



---

## Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval(a) samostatně a že jsem uvedl(a) veškeré použité informační zdroje v souladu s Metodickým pokynem o etické přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisů. V souladu s ust. § 46 odst. 6 tohoto zákona tímto uděluji nevýhradní oprávnění (licenci) k užití této mojí práce, a to včetně všech počítačových programů, jež jsou její součástí či přílohou, a veškeré jejich dokumentace (dále souhrnně jen „Dílo“), a to všem osobám, které si přejí Dílo užít. Tyto osoby jsou oprávněny Dílo užít jakýmkoli způsobem, který nesnižuje hodnotu Díla, a za jakýmkoli účelem (včetně užití k výdělečným účelům). Toto oprávnění je časově, teritoriálně i množstevně neomezené. Každá osoba, která využije výše uvedenou licenci, se však zavazuje udělit ke každému dílu, které vznikne (byť jen zčásti) na základě Díla, úpravou Díla, spojením Díla s jiným dílem, zařazením Díla do díla souborného či zpracováním Díla (včetně překladu), licenci alespoň ve výše uvedeném rozsahu a zároveň zpřístupnit zdrojový kód takového díla alespoň srovnatelným způsobem a ve srovnatelném rozsahu, jako je zpřístupněn zdrojový kód Díla.

V Praze dne 4. května 2018

.....

České vysoké učení technické v Praze

Fakulta informačních technologií

© 2018 Martin Melichar. Všechna práva vyhrazena.

*Tato práce vznikla jako školní dílo na Českém vysokém učení technickém v Praze, Fakultě informačních technologií. Práce je chráněna právními předpisy a mezinárodními úmluvami o právu autorském a právech souvisejících s právem autorským. K jejímu užití, s výjimkou bezúplatných zákonných licencí, je nezbytný souhlas autora.*

## **Odkaz na tuto práci**

Melichar, Martin. *Systém pro správu elektronických verzí literárních děl*. Bakalářská práce. Praha: České vysoké učení technické v Praze, Fakulta informačních technologií, 2018.



---

## Abstrakt

Cílem této bakalářské práce je ulehčit převod literárních děl do elektronické podoby. Výsledná aplikace předpokládá naskenování knihy a poté načtení výsledku do aplikace. Následně přichází na řadu správa knihy např. oprava chyb z naskenování nebo úprava její části. Vytvořené řešení umožňuje upravovat, mazat nebo přidávat dílčí součásti knihy např. autor práce. //todo nějaké čísla (zrychlení atd.)// Hlavním výsledkem je zrychlení vytváření a snadné upravování elektronických verzí literárních děl.

**Klíčová slova** webová aplikace, redakční systém, návrh a implementace, správa elektronických literárních děl, TEI, Slim, PHP

---

## Abstract

Sem doplňte ekvivalent abstraktu Vaší práce v angličtině.

**Keywords** web application, management system, design and implementation, digitalized literary works, TEI, Slim, PHP



---

# Obsah

Odkaz na tuto práci . . . . .	viii
<b>Úvod</b>	<b>1</b>
Cíl práce . . . . .	1
Úvod do problematiky . . . . .	2
<b>1 Analýza</b>	<b>3</b>
1.1 Existující řešení . . . . .	3
1.1.1 Emeditor . . . . .	3
1.1.2 Editix . . . . .	4
1.1.3 EditPad Pro . . . . .	5
1.1.4 Essential XML Editor . . . . .	5
1.1.5 Exchanger XML Editor . . . . .	6
1.1.6 Vyhodnocení . . . . .	6
1.2 Technologie . . . . .	7
1.2.1 Jazyk . . . . .	7
1.2.2 Databáze . . . . .	8
1.2.3 Elektronický formát . . . . .	9
1.3 Software . . . . .	11
1.3.1 Microframework vs framework . . . . .	11
<b>2 Návrh aplikace</b>	<b>13</b>
2.1 Wireframy . . . . .	13
2.1.1 Login . . . . .	14
2.1.2 Hlavní stránka . . . . .	14
2.1.3 Metadata . . . . .	14
2.1.4 Přílohy . . . . .	14
2.1.5 Autoři a vydavatelé . . . . .	14
2.1.6 Text . . . . .	15
2.2 Databázové schéma . . . . .	15

2.3	Autentizace . . . . .	17
<b>3</b>	<b>Implementace</b>	<b>19</b>
3.1	Import . . . . .	19
3.1.1	Stávající sbírka děl . . . . .	19
3.1.2	Přílohy . . . . .	21
3.2	Aplikace . . . . .	22
3.2.1	Kostra aplikace . . . . .	22
3.2.2	Login . . . . .	22
3.2.3	Seznam děl . . . . .	23
3.2.4	Seznam autorů a vydavatelů . . . . .	23
3.2.5	Metadata . . . . .	23
3.2.6	Přidat nového autora nebo vydavatele . . . . .	23
3.2.7	Upravit autora nebo vydavatele . . . . .	23
3.2.8	Přílohy . . . . .	23
<b>4</b>	<b>Testování</b>	<b>25</b>
4.1	Uživatelské testy . . . . .	25
4.2	Funkční testy . . . . .	25
<b>5</b>	<b>Budoucí rozšíření</b>	<b>27</b>
5.1	Nově vzniklá díla . . . . .	27
5.2	Náhled . . . . .	27
5.3	Zobrazení děl . . . . .	27
	<b>Závěr</b>	<b>29</b>
	<b>Literatura</b>	<b>31</b>
	<b>A Screenshoty</b>	<b>33</b>
	<b>B Seznam použitých zkratk</b>	<b>41</b>
	<b>C Obsah přiloženého CD</b>	<b>43</b>

---

## Seznam obrázků

2.1	Úprava autora . . . . .	15
2.2	Databázové schéma . . . . .	16
A.1	Přihlašování . . . . .	33
A.2	Ukázka podoby díla . . . . .	34
A.3	Hlavní stránka . . . . .	35
A.4	Metadata . . . . .	36
A.5	Přílohy . . . . .	37
A.6	Seznam autorů a vydavatelů . . . . .	38
A.7	Text díla . . . . .	39



---

# Úvod

Elektronická literární díla stále rozšiřují pole své působnosti, ať mluvíme o nakupování nebo zobrazování knih na počítači či o jejich snadném čtení v e-čtečkách. V Austrálii se rozdíl prodeje elektronických knížek v roce 2008 oproti roku 2009 rapidně zvýšil, a to o více než 100 %. Ve Spojených státech amerických se v lednu roku 2012 zvýšil prodej e-knih pro dospělé o 49,4 % a e-knih pro děti a mládež o 475,1 %, než tomu bylo v lednu 2011 [1]. Zde vidíme prudký nárůst prodeje e-knih, z toho můžeme usoudit, že rapidně roste oblíbenost elektronických děl. Zobrazení a následné čtení e-knih na našich elektronických zařízeních je nejenom snazší, ale i pohodlnější oproti zapůjčování či koupi papírových knih.

Výsledek této práce je přednostně určen pro zaměstnance Ústavu české literatury Akademie věd České republiky (UČL AV). Pracovníci ústavu získávají, ať formou darů či koupí, postupně více literárních děl. Tato díla následně naskenují a převedou do elektronické podoby. Nicméně často to bývají historická díla, a proto se stává, že naskenovaný výsledek přesně neodpovídá tištěné podobě knihy. Proto je nezbytné, aby pracovníci ústavu ručně opravovali chyby a měli možnost doplňovat chybějící části v elektronické podobě díla. Zde přichází na řadu aplikace, která bude umožňovat rychlejší a efektivnější úpravu nově vzniklého díla.

Na základě poznámek a připomínek zaměstnanců bude tato aplikace efektivní a upravená na míru. Toto téma jsem si zvolil, protože je velmi aktuální a výsledek by mohl výrazně ulehčit práci při vytváření literárních elektronických děl.

## Cíl práce

Prvním cílem je rešerše existujících aplikací pro tvorbu a správu literárních děl a dále porovnání a výběr nejvhodnějšího frameworku pro tvorbu aplikace. S tím úzce souvisí výběr samotného jazyka pro implementaci. Poté přijde na

řadu porovnání a volba nejvhodnějšího textového formátu pro e-knihy. Dalším krokem je shromáždění poznámek a připomínek od pracovníků pro vytvoření aplikace na míru.

Cílem praktické části práce je navrhnout a implementovat redakční systém pro správu a snadné vytváření elektronických verzí literárních děl, otestovat systém na reálných datech a uživatelích a řádně jej zdokumentovat. Systém bude umožňovat správu nejen samotných dokumentů, ale i souborů k nim přidruženým, zejména fotografií a obrázků. Ve výsledné aplikaci bude implementován jednoduchý filtr a snadné vyhledávání v seznamu sbírek. Aplikace bude umožňovat autentifikaci a bude zároveň poskytovat správu uživatelů a jejich rolí. V aplikaci bude administrátorovi umožněna registrace nového uživatele, který si následně upraví heslo podle své potřeby.

## Úvod do problematiky

V současnosti mají pracovníci UČL AV okolo 1700 literárních děl uložených v databázi ve formátu XML. Díla jsou rozdělena do dvou částí, jak je vidět na příkladu Vytržené listy A.2. První část je identifikována tagem hlavička a druhá tagem text. Hlavička obsahuje metadata jako je titul díla nebo rok vydání a ve druhé části je samotný text rozčleněn například do tagů sbírka a báseň.



---

# Analýza

## 1.1 Existující řešení

Před návrhem a implementací aplikace bylo potřeba řádně prozkoumat existující řešení problému, technologií k tvorbě a elektronické formáty, které se užívají pro uchování elektronických děl. Vzhledem k roztoucí poptávce po elektronických dílech na úkor papírových se podle očekávání objevilo mnoho aplikací pro editaci těchto děl.

Následují informace o pěti editorech, které uvádí server [2]:

- Emeditor,
- Editix,
- EditPad Pro,
- Essential XML Editor,
- Exchanger XML Editor.

### 1.1.1 Emeditor

Emeditor<sup>1</sup> patří podle [2] mezi nejlepší XML editory. Tento software je hlavním produktem americké firmy Emurasoft, Inc. sídlící v Redmondu ve Washingtonu. Firma se nadále stará o podporu i vývoj. Nicméně autorem editoru je Yutaka Emura. Samotná aplikace již vyhrála 24 mezinárodních cen v kategoriích nejlepší webový nástroj nebo nejlepší aplikace roku 2008.

#### Výhody:

- poslední release v17.5.0 vyšla 27. února 2018, je tedy pravidelně aktualizován,

---

<sup>1</sup>domovská stránka: <https://www.emeditor.com/>

## 1. ANALÝZA

---

- podpora velkých souborů,
- použití více jader při větší zátěži,
- kódování UTF-8,
- konfigurovatelná kontrola pravopisu.

### **Nevýhody:**

- aplikace je placená, ale nabízí trial verzi na 30 dní,
- existuje free verze, nicméně v ní chybí zásadní funkce,
- editor je pouze pro Windows,
- chybí přehledný průvodce základních funkcí po prvním spuštění.

### **1.1.2 Editix**

Editix<sup>2</sup> je produktem francouzské společnosti JAPISoft SARL. Editor vytvořil Alexandre Brillant. Systém je velice přehledný a intuitivní. Na trhu je systém od roku 2004 a nejnovější verze je EditiX XML Editor 2017 v15. Zákazníci, kteří využívají tento software jsou převážně vzdělávací instituty od University of Oxford po University of Arizona.

### **Výhody:**

- přehledný program,
- aplikace je pro platformy Windows, Linux, MacOS,
- existuje EditiX Community Edition, která je zadarmo,
- mnoho užitečných funkcí např. Find and Replace,
- obsahuje inteligentní našeptávač, který pomáhá uživatelům.

### **Nevýhody:**

- verze pro je placená, ale nabízí trial verzi na 30 dní,
- existuje lite verze, které chybí spoustu funkcí,
- zaplacení licence se vztahuje pouze na jednoho uživatele,
- první update je zdarma, další se musí zaplatit.

---

<sup>2</sup>domovská stránka: <http://www.editix.com/index.html/>

### 1.1.3 EditPad Pro

EditPad Pro<sup>3</sup> je výhradně textový editor, který lze použít například pro HTML, Javascript nebo XML. Software podporuje změnu jazyka například do francouzštiny, němčiny, polštiny nebo švédštiny. Projeck Just Great Software, pod kterým byla vyvinuta tato aplikace vznikl v roce 1996. Autorem projektu je Jan Goyvaerts, který je zároveň hlavním ředitelem vývojáři projektu.

**Výhody:**

- program je obecný textový editor,
- obarvená syntaxe,
- dobře pracuje s velkými soubory,
- podpora UTF-8,
- existuje EditPad Lite verze, která je zdarma.

**Nevýhody:**

- nemá explicitní podporu pro XML,
- aplikace je pouze pro Windows,
- chybí vyhledávání v souborech,
- verze Pro je placená, ale existuje verze Lite.

### 1.1.4 Essential XML Editor

Essential XML Editor<sup>4</sup> je jednoduchý XML editor. Jeho klíčovou vlastností je vestavěný XML validátor. Vývojáři dříve pojmenovali program Open XML Editor, ale po zavedení poplatku za některé funkce projekt přejmenovali. Autorem je Dieter Köhler.

**Výhody:**

- program pracuje jako textový editor,
- možnost rychle zjistit zda je soubor validní,
- trial verze není časově omezená,
- vstupní soubor může být v různém kódování,
- klávesová zkratka pro každý příkaz.

---

<sup>3</sup>domovská stránka: <http://www.editpadpro.com/>

<sup>4</sup>domovská stránka: <http://www.philo.de/xmledit/>

### **Nevýhody:**

- výstupní soubor pouze v UTF-8 kódování,
- aplikace je pouze pro Windows,
- pro zpřístupnění některých funkcí nutnost zakoupit klíč,
- poměrně zastaralý desing aplikace.

### **1.1.5 Exchanger XML Editor**

Exchanger XML Editor<sup>5</sup> je určen pro snadnou editaci, prohlížení, správu a konverzi XML souborů. Exchanger pomáhá svojí širokou nabídkou funkcí XML autorům a vývojářům. Software je produktem firmy Cladonia, která se zaměřuje na vývoj XML aplikací.

### **Výhody:**

- nabízí plnou verzi na 30 dní,
- aplikace je dostupná na všech platformách,
- možnost zobrazení základního náhledu,
- poskytuje podporu pro XML formou stáhnutí balíčku,
- automatická kontrola, jestli je soubor validní.

### **Nevýhody:**

- při instalaci nutnost najít cestu k JRE manuálně,
- zastaralý software,
- časově neomezenou verzi je nutno zakoupit,
- poslední update proběhl v roce 2010.

### **1.1.6 Vyhodnocení**

Z potenciálních konkurentů můžeme vyškrtnout Emeditor, protože je placený. Existuje neplacená verze Emeditoru, ale zároveň v ní chybí například možnost vložení tagu před a za označený text. Aplikace Editix je také zpoplatněna. Velkou nevýhodou je placení každého dalšího updatu editoru zvlášť. EditPad Pro je také placený a ve volně dostupné verzi není možnost vyhledávání v souborech. Navíc explicitně nepodporuje XML. Z názvu editoru Essential XML Editor lze odvodit, že je určený pro XML, ale ihned po spuštění aplikace je zřejmá dlouhodobá neaktualizace designu. Podobný problém s designem má Exchanger XML Editor, u kterého navíc poslední update softwaru proběhl v roce 2010.

---

<sup>5</sup> domovská stránka: <http://www.exchangerxml.com/editor/>

## 1.2 Technologie

Mezi první otázky patřilo, v jaké technologii se bude aplikace psát. Po domluvě s pracovníky UČL AV jsem měl vytvořit webovou aplikaci, která má mít přívětivé uživatelské rozhraní pro snadnou a rychlou správu elektronických děl.

Další důležitou otázkou bylo, ve kterém elektronickém formátu se budou díla tvořit a uchovávat. Existuje celá řada formátů, proto bylo pro budoucí možné rozšíření důležité domluvit se s pracovníky na správném formátu.

### 1.2.1 Jazyk

U jednoduchých webových aplikací stačí, když klientská strana pošle požadavek na serverovou část, ta jej vyhodnotí a pošle odpověď zpět. Pro implementaci tohoto případu se nejčastěji používá architektura client/server [3]. Jedna z výhod této architektury je, že klientská strana programu je oddělena od serverové části. Výpočty jsou prováděny na straně serveru, proto nejsou požadovány vysoké nároky na výpočetní techniku počítače, na kterém běží klientská část. Tato architektura také chrání data jejich uložením na server, což je jedna z nejlepších metod ochrany.

Nevýhodou může být vysoká cena zařízení pro provoz případně je nutné mít pro správu serveru systémového administrátora.

Srovnání jazyků a informace pro webové aplikace uvedené na [3].

- PHP – jeden z nejrozšířenějších jazyků, který podporuje většina poskytovatelů webhostingu. Tento jazyk je široce využíván mezi uživateli a obsahuje mnoho standardních knihovních funkcí. Hodí se pro malé nebo středně velké projekty.
- Ruby – navzdory tomu, že Ruby je mladý jazyk, těší se velké oblíbenosti mezi webovými vývojáři. Nejznámější framework Ruby on Rails umožňuje rychle vytvářet vzorové nebo malé projekty. Pro nové uživatele je z důvodu daných pravidel Ruby on Rails poměrně složitý. Neznalý uživatel bude mít na začátku problém i s implementací jednoduchých funkcí.
- Python – díky nástupu velkých frameworků - například Django, je možné použít Python ve velkých webových projektech. Syntaxe kódu je podobná Ruby, ale hlavní rozdíl činí ideologie psaní kódu. Nicméně narozdíl od Ruby se Python rozvíjí rychleji.
- C# ASP.NET – pro správné fungování jazyka je třeba zařídit ISS server, který je nezbytný pro mnoho komerčních projektů od firmy Microsoft. Jiným řešením může být použití serveru Mono, který ale není stabilní a může obsahovat mnoho chyb.

- JAVA – tento široce využívaný jazyk lze aplikovat na velké či korporální projekty. Je poměrně rychlý a obsahuje spoustu již vyřešených složitých problémů, které nemusí uživatel znovu řešit. Pro začátek je vyžadováno poměrně velké množství znalostí, aby fungovala základní kostra programu, což se považuje za významnou nevýhodu.

Pro tento projekt postačí menší a jednodušší aplikace, kterou zaměstnanci UČL AV snadno zprovozní na svých serverech. Pracovníci nemají k dispozici server ISS, proto vypadává z výběru C# ASP.NET. S Ruby nemám žádné zkušenosti a Python jsem dříve používal jen okrajově. PHP i JAVA jsou velice rozšířené jazyky ve webovém inženýrství, nicméně k PHP mám kladnější vztah. Znalosti PHP jsem využil hned v několika předmětech při studiu. Na základě této analýzy byl vybrán jazyk PHP.

### 1.2.2 Databáze

Srovnání několika open-source databázových řešení podle [4].

- CUBRID<sup>6</sup> – byl navržen a optimalizován pro webové aplikace. Hodí se pro práci s velkými daty, nebo při nutnosti použití více dotazů najednou. Výhodou je možnost online zálohy a grafické rozhraní pro jazyky PHP, Python, Perl a Ruby. Na druhou stranu manuál existuje pouze v angličtině nebo korejštině.
- Firebird<sup>7</sup> – je prvním zástupcem relační databáze. Systém funguje již od roku 1981, který lze použít na Linuxu, Windows a dalších Unixových platformách. Firebird má obsáhlou komunitu uživatelů. V databázi lze využít mnoho vývojářských nástrojů a lze se do ní přihlásit pomocí účtu na microsoftu. Naopak chybí integrace mezi ostatními databázovými systémy.
- MariaDB<sup>8</sup> – byla vytvořena původními vývojáři MySQL. MariaDB využívají dnes největší společnosti jako jsou Google nebo Facebook. Ochrana dat je na špičkové úrovni. Systém působí na trhu přes 20 let. Jako nevýhoda se dá počítat chybějící rozhraní pro mezipaměť.
- MongoDB<sup>9</sup> – byla vytvořena v roce 2007 jako řešení pro větší projekty. Díky svým sponzorům a podporovatelům si tato databáze uchovává myšlenku být jednoduchá a efektivní. Dokáže zpracovat poměrně rychle velké množství dat. Po zaznamenání systémové chyby následuje rychlé obnovení dat.

---

<sup>6</sup>domovská stránka: <https://www.cubrid.org/>

<sup>7</sup>domovská stránka: <http://www.firebirdsql.org/en/start/>

<sup>8</sup>domovská stránka: <https://mariadb.org/>

<sup>9</sup>domovská stránka: <https://www.mongodb.com/>

- MySQL<sup>10</sup> – funguje již od roku 1995. Dnes je využívána jako standardní databáze pro menší i větší projekty. Běží na všech známých operačních systémech a funguje i při výpadku internetu. Velkou výhodou je oddělený server od vývojového prostředí. Naopak nevýhodou je delší prodleva mezi novými aktualizacemi.
- PostgreSQL<sup>11</sup> – má za sebou 15 let aktivního vývoje a patří mezi databáze, které běží na všech hlavních operačních systémech. S použitím PostgreSQL může uživatel vytvořit vlastní metody nebo nestandardní datové typy. Mnoho vřetavěných procedur lze spouštět pomocí mnoha programovacích jazyků, jako je Java, Perl, Python nebo C/C++. Vývoj je řízen pouze komunitou.
- SQLite<sup>12</sup> – se podle [4] považuje za nejrozvinutější databázi na světě. Vývoj začal v roce 2000 a používali jej významné firmy jako je Facebook, Apple nebo Microsoft. Po každé aktualizaci, vývojáři zveřejní podrobný výpis změn. K dispozici je kvalitní podpora a knihovna, která na úkor velikosti paměti pracuje velmi rychle. SQLite se nedoporučuje pro rozsáhlé webové aplikace a velké množství dat.

Výsledná aplikace bude obsahovat méně než 10 tabulek. Aplikace využije jednoduché dotazy a pro chod aplikace nebude potřeba posílání několika dotazů současně. Do budoucna se počítá s rozšířením aplikace, proto je výhodou, když je databáze udržována a aktualizována. Na základě analýzy a protože tento projekt bude patřit spíše k těm menším, byla vybrána databáze SQLite.

### 1.2.3 Elektronický formát

*„Mezi formáty, v nichž můžeme číst elektronickou literaturu, jsou jednak ty, které byly pro tento účel přímo vytvořené, ale také ty, v nichž se e-knihy publikovaly prostě proto, že nebylo mnoho jiných alternativ. Toto se týká zejména stavu v 90. letech, kdy vznikaly kopie (převážně papírových) knih převedené do formátů jako jsou TXT, HTML či RTF.“*[5]

Elektronickému formátu je třeba věnovat zvláštní pozornost, protože se do budoucna počítá s rozšířením aplikace o část, která se bude věnovat zobrazením těchto děl pro veřejnost.

Seznam vybraných formátů pro e-knihy uvedené v [5].

- Archos Diffusion – je formát vytvořen francouzskou firmou ArchosDiffusion. Koncovka názvů souborů je .aeh. Formát byl vytvořen pro uchovávání literatury v elektronické podobě a patří do skupiny formátů založených na XML. Otevírat soubory lze v programu Archos Player nebo

<sup>10</sup> domovská stránka: <https://www.mysql.com/>

<sup>11</sup> domovská stránka: <https://www.postgresql.org/>

<sup>12</sup> domovská stránka: <https://www.sqlite.org/index.html>

ve volně dostupné aplikaci Visual Vision EbooksReader. Postupem času uvádá zájem o tento formát.

- AZW – formát vyvinuli vývojáři Amazonu a používá koncovku .azw. Tento formát byl v zhotoven pro uchování elektronických děl v internetovém knihkupectví společnosti Amazon. Knihy lze číst ve všech dostupných verzích čteček Kindle. Díky oblíbenosti čteček Kindle patří AZW mezi nejrozšířenější formáty na světě. Velkou nevýhodou je jeho uzavřenost, protože knihy ve formátu AZW prakticky nelze číst v jiných čtečkách než jsou Kindle.
- EPUB – patří mezi nejpopulárnější formáty na světě i v České republice. Vytvořilo jej sdružení International Digital Publishing Forum a je založený na XML. Knihy ve formátu EPUB lze číst na většině čtecích zařízeních s výjimkou čteček Kindle, nicméně existuje možnost převodu formátu EPUB do jiného, který dokáže číst i čtečky Kindle. Tento otevřený formát podporuje Digital rights management (DRM) ochranu a proto si získal oblibu i u nakladatelů.
- Hypertext Markup Language (HTML) – se primárně používá pro tvorbu webových stránek. V 90. letech, kdy se začaly poprvé vytvářet elektronické podoby knih, nebyl ještě vyvinut žádný formát pro jejich zobrazení, a proto se z nutnosti používal také HTML. Převod probíhal nejprve naskenováním díla a poté nahráním do aplikace Optical Character Recognition (OCR). Výstupem aplikace bylo dílo ve formátu HTML. Ke čtení postačil webový prohlížeč. Nicméně s nástupem formátů navržených pro elektronické knihy se přestal HTML používat.
- Portable Document Format (PDF) – vyvinutý společností Adobe v roce 1993 byl primárně určen pro uchování souborů pro tisk. Patří mezi formáty, které nebyly vytvořeny pro elektronickou literaturu, ale narodil od ostatních, se tak používá dodnes. Velkou předností je nezávislost na platformě, protože s PDF lze pracovat téměř na všech operačních systémech.
- Plain text – patří mezi formáty, které nebyly určeny pro uchovávání e-knih. V 90. letech se používal pro zobrazení elektronické literatury, protože nebylo tolik jiných možností. Díky malé datové velikosti souborů a možnosti čtení souborů neomezeně na platformě se stále používá. TXT nedovoluje formátování a nepodporuje vložení obrázků, videí či zvukových stop.
- Text Encoding Initiative (TEI) – byl vytvořen TEI konsorciem primárně pro elektronickou literaturu. Formát je využíván ve výukových projektech i knihovnách po celém světě. Řadí se do skupiny formátů, které jsou založeny na XML. TEI se označuje jako nastavitelný, protože uživatel



může podle své vůle přidat, předefinovat nebo přejmenovat tagy a jejich atributy. Formát se dá použít pro různé druhy textů.

Drtivá většina děl dostupná pracovníkům UČL AV je právě ve formátu XML. Avšak tento formát není primárně určen pro uchovávání elektronických verzí literárních děl, proto byl pro mou práci po diskuzi s pracovníky UČL AV vybrán textový formát TEI.

## 1.3 Software

### 1.3.1 Microframework vs framework

V následujícím textu vycházím z [6].

Framework poskytuje skoro vše, co programátor potřebuje, od obsluhy webových požadavků po komunikaci s databází. Obsahuje i komponenty, které vývojář nemusí nikdy použít, nicméně z hlediska rozšiřitelnosti jsou výhodné.

Microframework je označením pro framework, který obsahuje pouze nejnutnější komponenty k vývoji webové aplikace. Microframeworky bývají přizpůsobeny menším aplikacím nebo aplikacím s konkrétním účelem. Pro rozšíření funkčnosti je potřeba přidat dané komponenty.

Reálně je microframework sbírka nejnutnějších potřebných komponent pro potřebu webových aplikací. Obsluha microframeworku dostane HTTP požadavek, který zpracuje daný kontroler a jen pošle odpověď obvykle ve formátu HTML zpět. Některé microframeworky obsahují další nástroje pro manipulaci s HTTP požadavky. Mnoho vývojářů používá raději velké frameworky, jako jsou Laravel nebo Symfony. Tyto frameworky disponují spoustou již vyřešených problémů a mají velkou programátorskou základnu. Nicméně potřebují více času pro pochopení a porozumění prostředí.

Jsou projekty, u kterých se vyplatí použít velké frameworky, ale v případě projektů, kde není potřeba tolik funkcí, je vhodné použít microframework. Tento projekt se charakterizuje spíše menší náročností a nebude obsahovat složité požadavky. Aplikace nebude mít mnoho stránek a proto ideláním řešením bude microframework.

5 nejlepších microframeworků podle [7].

- Slim<sup>13</sup> – je považován za jeden z nejlepších PHP microframeworků. Umožňuje snadno vytvořit kvalitní webovou aplikaci. Díky nastavitelné a modulární architektuře poskytuje vývojářům přesně to, co potřebují. Slim dovoluje vkládat závislosti, proto jej lze použít společně s externími nástroji.
- Silex<sup>14</sup> – byl vyvinut z frameworku Symfony, aby byl co nejmenší a zároveň poskytoval základní funkčnost. Nakonec vznikly dvě verze. Fat

---

<sup>13</sup>domovská stránka: <https://www.slimframework.com/>

<sup>14</sup>domovská stránka: <https://silex.symfony.com/>

## 1. ANALÝZA

---

verze má v sobě komponenty ze Symfony, Twig, šablonovací systém a jiné. Druhá slim verze obsahuje základní systém routování a několik procedur.

- Wave<sup>15</sup> – používá architekturu MVC. Neobsahuje doplňkové knihovny a klade důraz na rychlost a optimalizaci. Wave podporuje Apache i Nginx servery.
- Limonade<sup>16</sup> – se zaměřuje obdobně jako Wave na jednoduchost. Limonade je velmi rychlý a snadný ve vývoji aplikací. Nicméně je až extrémně malý a nelze jej rozšířit o složitější funkčnost. Spoléhá pouze na globální funkce.
- Lumen<sup>17</sup> – je microframework odvozený z asi nejrozšířenějšího PHP frameworku Laravel. Pokud si programátor není jistý velikostí svého projektu, je výhodné použít Lumen, protože stačí veškerý kód z Lumenu převést do Laravelu a vše bude fungovat jak má.

Po prostudování a podrobné analýze současně dostupných microframeworků byl vybrán Slim. Zároveň bude použit šablonovací systém TWIG, který Slim podporuje.

---

<sup>15</sup>domovská stránka: <https://www.waveframework.com/>

<sup>16</sup>domovská stránka: <https://limonade-php.github.io/>

<sup>17</sup>domovská stránka: <https://lumen.laravel.com/>

## Návrh aplikace

Návrh aplikace je velice důležitou fází projektu, ve které by mělo dojít ke sjednocení požadavků zadavatelů a reálného provedení. Vytvořením dobrého návrhu se zamezí případným kolizím a zároveň proběhne první interakce mezi zákazníkem a dodavatelem. Ze strany UČL byl vznesen požadavek, aby v aplikaci byl použit Bootstrap<sup>18</sup>, protože jsou na něj zvyklí. Databázové schéma bylo čistě na mém rozhodnutí.

### 2.1 Wireframy

Jedním z nejdůležitějších úkolů bylo navrhnout, jak budou jednotlivé stránky vypadat a kolik jich aplikace bude obsahovat. Základní rozložení stránek bylo na mém rozhodnutí. Design měl být jednoduchý, uživatelsky přívětivý a moderní. Tyto návrhy musely být a byly schváleny pracovníky UČL AV.

Grafika měla být jednoduchá, bez složitých a obsáhlých prvků. Pro snadnější a moderní stylování aplikace byl použit Bootstrap. Jedná se o volně dostupnou knihovnu, která dovoluje stylovat vzhled aplikace pouze přidáním určitých tříd k elementům v HTML. Při dodržování pravidel Bootstrapu není potřeba vkládat obsáhlé CSS styly. Pomocí Bootstrapu lze poměrně snadno vytvořit responzivní aplikaci. Toho by se dalo využít v případě budoucím rozšíření o mobilní zobrazení aplikace.

Pro návrh GUI jsem se rozhodl použít webovou aplikaci Moqups<sup>19</sup>. Její bezplatná verze nabízí velké množství šablon pro grafické návrhy. Od základních prvků jako je tlačítko nebo nadpis, po mírně složitější formuláře. Velikou výhodou je, že Moqups má v nabídce prvky Bootstrapu.

<sup>18</sup> domovská stránka: <https://v4-alpha.getbootstrap.com/>

<sup>19</sup> domovská stránka: <https://moqups.com/>

### 2.1.1 Login

Pro úvodní přihlašovací stránku jsem vybral jednoduchý formulář, ve kterém je email a heslo, viz obrázek A.1. Pozadí, které by se dalo v budoucím rozšíření měnit, je jako na všech ostatních stránkách bílé.

### 2.1.2 Hlavní stránka

Po úspěšném přihlášení do aplikace se zobrazí hlavní stránka, která je rozdělena do tří částí, jak je vidět na obrázku A.3. V hlavičce se nachází společně s nadpisem odkaz na seznam autorů a vydavatelů děl a dropdown s možnostmi uživatele. Dále je zde filtr aplikovatelný na seznam zobrazených děl. Požadavek od UČL AV byla možnost filtrovat elektronickou literaturu podle autora, roku vydání, textu obsaženého v díle a statusu. Třetí část obsahuje samotný seznam děl. Tento seznam je ve tvaru tabulky se sloupci název díla, autor, rok vydání, status, odkaz na přílohy a nezbytné akce a tabulka se dá seřadit podle sloupců. U seznamu děl lze nastavit počet zobrazených děl a obsahuje rychlý vyhledávač textu.

### 2.1.3 Metadata

Na obrázku A.4 je vidět rozložení stránky pro úpravu základních údajů o dílu. Opět se skládá ze tří částí. První obsahuje navigaci, dropdown pro přihlášeného uživatele a nadpis. Druhá část je věnována autorům a vydavatelům. V tomto oddílu bude umožněno přidávat a odebírat autora nebo vydavatele. Poslední sekce obsahuje formulář pro úpravu metadat literárního díla. Společně s pracovníky UČL AV byly vybrány atributy, které přímo nesouvisí s obsahem díla, nýbrž popisují samotnou publikaci.

### 2.1.4 Přílohy

Součástí zadání práce je správa příloh, zejména scanů stránek. Přílohám se věnuje právě tento segment. Obsahem wireframu A.5 jsou naskenované jednotlivé stránky daného literárního díla. V horní části se opět objevuje navigace, uživatelské funkce a nadpis. Následuje sekce věnovaná hromadnému uploadu skenů do níže zobrazené fotogalerie.

### 2.1.5 Autoři a vydavatelé

Pro správu autorů a vydavatelů byl použit návrh zobrazený v příloze A.6. V horní části se vyskytuje navigace, nezbytné funkce a nadpis. Dále je zde umístěna tabulka záznamů, ve kterých lze snadno vyhledávat. Každý záznam lze upravit a smazat. Pro upravení záznamu byl navržen wireframe 2.1, který obsahuje navigaci, funkce pro uživatele a zjednodušený formulář. Formuláře pro upravení vydavatele nebo autora jsou totožné.

Seznam děl List

user@email.com ▼

Udaje o: Alois

Jméno  
Alois

Příjmení  
Zadejte příjmení

Korporace  
Zadejte korporaci

Uložit

Obrázek 2.1: Úprava autora

### 2.1.6 Text

Nejdůležitějším prvkem aplikace byla bezesporu možnost upravovat text literárního díla. Pro tuto funkci byla navržena stránka A.7. Jako u předešlých wireframů, horní část se věnuje navigaci, funkcím uživatele a nadpisu. Dále je vidět rozložení stránky na dva oddíly. Jeden pro text knihy a druhý pro možnost vkládání značek. Menu vpravo bylo navrženo vzhledem k nadměrné velikosti první části tak, aby bylo viditelné, když se uživatel posune níž.

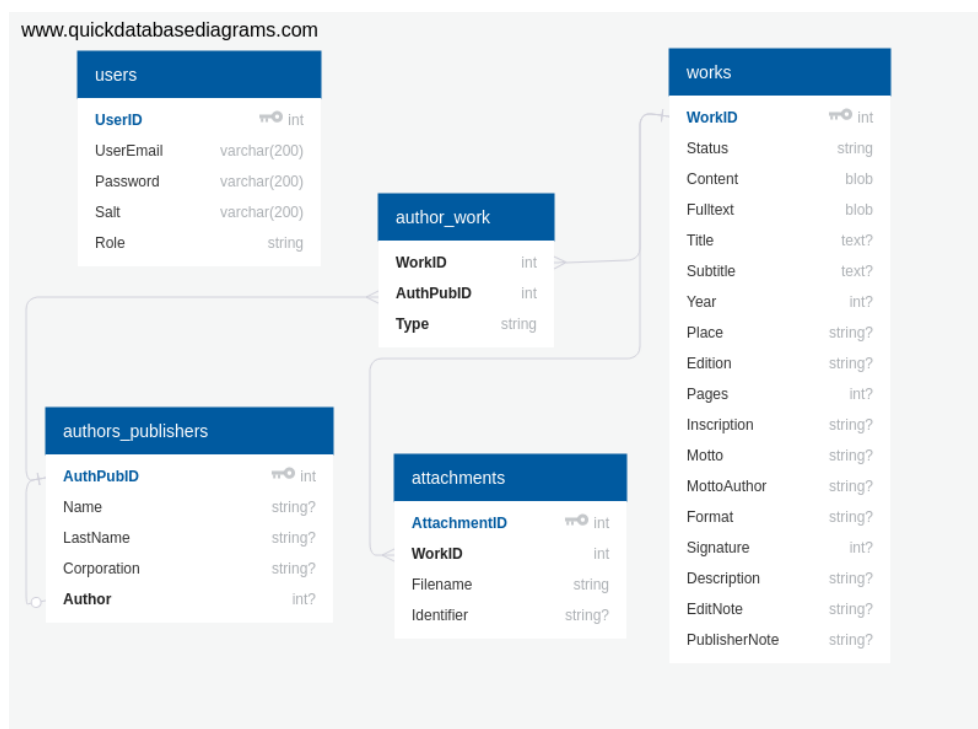
## 2.2 Databázové schéma

Neméně důležitou součástí projektu byla databáze. Ze strany UČL AV nebyly vzneseny žádné požadavky na podobu databáze, proto se mohlo schéma přizpůsobit dle potřeby aplikace.

Pro model databáze byla použita webová aplikace<sup>20</sup>. Schéma je navrženo minimalisticky, ale tak aby zároveň pokrývalo veškeré potřeby aplikace. Výsledný návrh databáze na obrázku 2.2 obsahuje 5 tabulek s jednoduchými vazbami.

<sup>20</sup> domovská stránka: <https://www.quickdatabasediagrams.com/>

## 2. NÁVRH APLIKACE



Obrázek 2.2: Databázové schéma

V tabulce uživatelů je uloženo UserID, UserEmail a heslo rozdělené do dvou sloupečků viz sekce 2.3. Uživatelské role jsou identifikovány pomocí sloupce Role.

Důležitá tabulka je **author\_work**. Obsahuje reference na tabulky **works** a **authors\_publishers**. Třetím atributem je typ spojení. Tento typ určuje, jestli je spojen autor nebo vydavatel díla. Každý záznam v tabulce propojuje dílo a autora nebo vydavatele. Primárním klíčem je celá trojice.

AuthPubID, Name, LastName, Corporation a Author jsou atributy tabulky **authors\_publishers**. Záznamem může být reálná osoba, její pseudonym nebo společnost. V případě pseudonymu odkazuje záznam na reálnou osobu. Tuto referenci obsahuje sloupec Author, ve kterém je AuthPubID reálné osoby zaznamenané v **authors\_publishers** nebo je prázdný.

Nejobsáhlejší tabulka se nazývá **works**. Obsahuje všechny základní informace o dílu: název díla (Title), podtitul (Subtitle), datum publikace (Year), místo vydání (Place), pořadí vydání (Edition), počet stran sbírky (Pages), věnování autora (Inscription), motto (Motto), autor motto (MottoAuthor), formát naskenovaného díla (Format), podpis (Signature), popis díla (Description) a poznámka k vydání (EditNote). Ve **works** je atribut Status, který indikuje momentální stav díla. Status může nabývat hodnot domluvených s UČL AV Nové, Rozpracováno, Zkontrolováno a Hotovo. Celý text včetně tagů

je ve sloupci Content. Dále tabulka works obsahuje sloupec Fulltext, ve kterém je text díla bez tagů.

Poslední tabulka attachments je pro přílohy přidružené k dílům. Obsahuje AttachmentID, referenci na dílo (WorkID), Filename a poznámku k obrázku (Identifier).

## 2.3 Autentizace

Součástí zadání je požadavek na možnost autentizace. Ze strany ústavu nebyl vznesen žádný požadavek na ochranu hesel, proto jsem si způsob zabezpečení mohl vybrat. Heslo je zašifrované pomocí sha256. Pro větší bezpečnost hesla jsem přidal metodou solení hesla tzn., že se k uživatelskému heslu přidává náhodný text.





## Implementace

Při vytváření kódu bylo nutné dodržovat domluvené návrhy. Kód musel být přehledný pro případné rozšíření aplikace jiným vývojářem. Díky vybranému softwaru a jeho přívětivé příručce byla poměrně rychle vytvořena aplikace v základní podobě. Slim má obrovskou komunitu uživatelů, což je velice nápomocné při hledání řešení problému.

### 3.1 Import

Ještě před implementací se musely do databáze aplikace zpracovat sbírky z UČL AV. Podle očekávání byla data dodána ve formátu XML. Import nedoprovázely žádné komplikace zejména díky snadné ovladatelnosti a standardním knihovnám scriptovacího jazyka Python<sup>21</sup>. Současně s importem sbírek proběhlo zpracování příloh k nim přidružených.

#### 3.1.1 Stávající sbírka děl

Stávající díla mají pracovníci z UČL AV v databázi. Část z této sbírky děl je ústavem již zpracována, na zbylé části se stále pracuje. Dříve byla tato díla upravována pomocí kancelářského balíčku Microsoft Word<sup>22</sup>. Word je skvělý nástroj pro editaci jednotlivých souborů, nicméně nedovoluje spravovat soubory jako celek. V současné době pracovníci UČL AV čekají na novou aplikaci, protože jim chybí vhodný nástroj k úpravě děl.

Import literárních děl proběhl otevřením každého souboru a zpracováním skriptem. Skript získal informace o dílu, vytvořil nové nebo navázal na již existující autory či vydavatele a vytvořil záznamy v tabulkách. Pro práci s XML byla použita standardní knihovna ElementTree<sup>23</sup>. Python má stejnojmennou knihovnu pro práci s databází sqlite.

<sup>21</sup> domovská stránka: <https://www.python.org/>

<sup>22</sup> domovská stránka: <https://products.office.com/cs-cz/word>

<sup>23</sup> domovská stránka: <https://docs.python.org/2/>

### 3. IMPLEMENTACE

---

V ElementTree se na začátku zavolá funkce `getroot`. Tato funkce inicializuje proměnnou, ve které je obsah souboru a dá se v ní snadno přistupovat k jednotlivým tagům. Tag `hlavicka` indikuje údaje o dílu a obsah díla patří pod tag `text`. Funkce `find` slouží k nalezení a vrácení tagu podle jeho názvu.

Fulltextové vyhledávání v aplikaci vyžaduje text bez tagů. K tomu slouží funkce `itertext`, která ignoruje tagy a vrátí celý text. Naopak funkce pro získání obsahu včetně tagů není v ElementTree vestavěna. Nicméně potřebný text lze dostat spojením jiných funkcí. Obsah této funkce byl inspirován [8].

Některá díla mají více autorů. V takovém případě všechna jména zapsána v tagu `author`. Tato jména jsou ohraničena množinovými závorkami a oddělena středníkem. Mnoho autorů si v průběhu své tvorby vymyslelo pseudonym. Pseudonym autora je indikován znakem „=“. Vlevo od znaku je pseudonym a vpravo je pravé jméno autora. Funkce `doAuthors` vrací id autora nebo pole id autorů v závislosti na vstupním parametru `authors`. Tento parametr je typu string a pokud začíná znakem „{“, má dílo více autorů. Pro tento případ funguje funkce jako rekurze. Druhým parametrem funkce je indikátor rekurze (`recursion`).

Pro ilustraci je uveden začátek funkce zpracování autorů díla.

```
def doAuthors(authors, recursion):
    #there is no author
    if (authors == 'neznamy'):
        return -1
    #there are more authors
    if (authors[0] == '{'):
        authors = authors.replace('{', '')
        authors = authors.replace('}', '')
        authArr = authors.split(';')
        authArr[0] = ' ' + authArr[0]
        out = []
        for authorName in authArr:
            out.append(doAuthors(authorName, 1))
        return out
    #there is only author
    elif (authors.find('=') == -1):
        authors = authors.replace('(', '')
        authors = authors.replace(')', '')
        lastName, sep, name = authors.partition(',')
        name = name[1:]
        if (recursion == 1):
            lastName = lastName[1:]
        #create or find and select author id from DB
        return getAuthorId(name, lastName)
```

Pro obsluhu vydavatelů byla použita funkce `doPublisher`, která pracuje podobně jako funkce pro autory, nicméně se zde nevyskytují pseudonymy.

Import do tabulky `works` proběhl ve funkci `doWorks`. Funkce má v parametrech název díla (`title`), rok vydání (`year`), status, proměnnou, která obsahuje ostatní informace (`meta`), `fulltext`, text (`content`) a poznámku k vydání (`note`). Funkce vrací id vytvořeného díla. Příklad zjištění počtu stránek díla z parametru `meta`:

```
pages = ''.join(meta.find('stran').itertext())
```

Záznamů do tabulky `connection` byly vloženy pomocí funkce `doConnection`. Tato funkce má parametry `workId`, `indexId` a `typeOfConn`. První je id vytvořeného díla a `indexId` je pole id autorů nebo vydavatelů, které může obsahovat jeden prvek. `TypeOfConn` je typ propojení, kterým je buď `author` nebo `publisher`. Obsah funkce je pouze databázový dotaz typu `insert`.

### 3.1.2 Přílohy

Zaměstnanci UČL AV dodali společně s některými díly i jejich naskenované stránky. Adresář `scan` obsahoval podadresáře `img`, kde byly hlavní obrázky a `thumbs`, kde se nacházely zmenšené obrázky. Jednotlivá díla jsou nazvána číslem například `0001.xml`. Tato čísla byla zároveň adresářem v `img` i `thumbs`. Podle čísla se zjistilo, ke kterému dílu obrázky patří. V číselných podadresářích byly uloženy obrázky ve formátu `jpg` a u hlavních obrázků byl soubor `pages.csv`, ve kterém byly informace o jednotlivých stránkách. Přílohy byly uloženy s číselným názvem. Příklad záznamu v souboru `pages`: `002.jpg;Strana [1]`. Tuto informaci bylo nutné zpracovat a vložit do tabulky jako poznámku k příloze. Problém nastal při otevírání souboru `pages.csv`, protože pracovníci si nepamatovali v jakém kódování soubor uložili. Po chvilkovém bádání se zjistilo, že použili `cp1250`.

V aplikaci se přílohy přidávají do adresáře `images` a podadresáře nazvaném podle id díla. Aby bylo id tvořeno pěti číslicemi byly k němu zleva přidány 4 nuly. Názvy příloh jsou u každého díla tvořena číslem začínajícím od 1. Pro snadnější řazení se jméno skládá ze tří znaků, čísla a nezbytné nuly na začátku. Zmenšené obrázky mají ke jménu přidanou příponu `_small`, příklad názvu obrázku: `010_small.jpg`. Vytvářet adresáře lze pomocí knihovny `os`, čtení souboru zajistí knihovna `pandas` a obsluhu souborového systému obstará knihovna `shutil`. Pro import příloh byla použita následující funkce:

```
doAttachments
def doAttachments(workID, oldID):
    #create directory if not exists
    os.makedirs('./images/' + '{0:0=5d}'.format(workID),
                exist_ok=True)
    tmp = './scan/scan/img/'+str(oldID)+'/'
    tmpSmall = './scan/scan/thumbs/'+str(oldID)+'/'
    khe = pd.read_csv(tmp + 'pages.csv',
                     encoding='cp1250',
                     sep=';', header=None)
    khe.columns = ['id', 'poznámka']
    l = 1
    for value in khe.id.keys():
        db.execute(insertAttachmentSQL,
                   [workID, khe.poznámka[value],
                    str('{0:0=3d}'.format(l)) + '.jpg'])
        sh.copy(tmp + khe.id[value],
                './images/' +
                str('{0:0=5d}'.format(workID)) + '/' +
                str('{0:0=3d}'.format(l)) + '.jpg')
        sh.copy(tmpSmall + khe.id[value],
                './images/' +
                str('{0:0=5d}'.format(workID)) + '/' +
                str('{0:0=3d}'.format(l)) + '_small.jpg')
        l = l + 1
```

## 3.2 Aplikace

### 3.2.1 Kostra aplikace

základní aplikace pomocí slimframeworku

### 3.2.2 Login

popis login funkcí

**3.2.3 Seznam děl**

**3.2.4 Seznam autorů a vydavatelů**

**3.2.5 Metadata**

**3.2.6 Přidat nového autora nebo vydavatele**

**3.2.7 Upravit autora nebo vydavatele**

**3.2.8 Přílohy**



# Testování

testování aplikace

## 4.1 Uživatelské testy

## 4.2 Funkční testy





---

## Budoucí rozšíření

### 5.1 Nově vzniklá díla

Pracovníci UČL AV mají připravený nově oskenované díla, které chtějí aplikovat na v budoucnosti se předpokládá využití aplikace na nově naskenované díla.

### 5.2 Náhled

rychlý náhled díla při aktualizaci

### 5.3 Zobrazení děl

příklad pro další bakalářku, kde se bude čistě zajímat o frontend



---

## **Závěr**



---

## Literatura


- [1] D'Ambra, J.; Wilson, C. S.; Akter, S.: Application of the task-technology fit model to structure and evaluate the adoption of E-books by Academics. *Journal of the Association for Information Science and Technology*, ročník 64, č. 1, 2013: s. 48–64.
- [2] TEI Community: Editors [online]. [cit. 2018-03-22]. Dostupné z: <https://wiki.tei-c.org/index.php/Editors>
- [3] Parkhomenko, A.; Sokolyanskii, A.; Gladkova, O.; aj.: Investigation of remote lab design technologies. In *2015 XI International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH)*, Sept 2015, s. 92–95, [cit. 2018-03-31].
- [4] Jennifer Champagne: The Top 7 Free and Open Source Database Software Solutions [online]. [cit. 2018-04-3]. Dostupné z: <https://blog.capterra.com/free-database-software/>
- [5] Pokorný Lukáš: *Formáty elektronických knih: specifika a popularita*. Inflow: information journal [online], 2012, [cit. 2018-04-5]. Dostupné z: <http://www.inflow.cz/formaty-elektronickych-knih-jejich-specifika-popularita>
- [6] VLADIMIR KAZANKOV: FULL-STACK FRAMEWORK OR MICROFRAMEWORK, LARAVEL OR LUMEN? [online]. [cit. 2018-04-12]. Dostupné z: <https://belitsoft.com/laravel-development-services/full-stack-framework-or-microframework-laravel-or-lumen>
- [7] Ankur Kumar: Best 5 Lightweight PHP Frameworks for REST APIs Development [online]. [cit. 2018-04-12]. Dostupné z: <http://findnerd.com/list/view/Best-5-Lightweight-PHP-Frameworks-for-REST-APIs-Development/33833/>

## LITERATURA

---

- [8] Mark Amery: How do I get the full XML or HTML content of an element using ElementTree? [online]. [cit. 2018-04-29]. Dostupné z: <https://stackoverflow.com/questions/380603/how-do-i-get-the-full-xml-or-html-content-of-an-element-using-elementtree>

## Screenshoty



The screenshot shows a login form with two input fields and a button. The first field is labeled 'Email' and contains the placeholder text 'Email'. The second field is labeled 'Heslo' and contains the placeholder text 'Heslo'. Below these fields is a blue button with the text 'Login'.

Obrázek A.1: Přihlašování

```
▼<dilo>
  ▼<hlavicka>
    <autor>Babánek, Karel</autor>
    <titul>Vytržené listy</titul>
    <podtitul/>
    <misto>Praha</misto>
    ►<vydavatel>...</vydavatel>
    <rok>1896</rok>
    <vydani>[1.]</vydani>
    <stran>56</stran>
    <venovani/>
    <moto/>
    <autormota/>
    <format>159x119 mm</format>
    ►<popis>...</popis>
    <zdroj-signatura>ÚČL AV ČR; 246 VIII 5</zdroj-signatura>
    ►<edicnipoznamka>...</edicnipoznamka>
    ►<komentare>...</komentare>
  </hlavicka>
▼<text>
  ▼<spisovatel>
    <b>Karel Babánek</b>
  </spisovatel>
  <br/>
  ▼<sbirka id="1">
    ►<nadpis>...</nadpis>
    <br/>
    <strana id="1" netistena="ano">[1]</strana>
    ►<tiraz>...</tiraz>
    <br/>
    <strana id="2" netistena="ano">[2]</strana>
  ▼<oddil id="1">
    ▼<nadpis>
      <b>Z ULICE</b>
    </nadpis>
    <br/>
    <strana id="3" netistena="ano">[3]</strana>
  ▼<basen id="1">
    ►<nadpis>...</nadpis>
    <br/>
    ▼<strofa id="1">
      <v id="1">Jde krokem loudavým a ruce v zad</v>
      <v id="2">má skříženy a tupě hledí v před,</v>
      <v id="3">když v šeré ulice tmy závoj pad',</v>
      <v id="4">a plynu žlutavý plá v chodník svit.</v>
    </strofa>
```

Obrázek A.2: Ukázka podoby díla



List

user@email.com ▲

Změna hesla

Něco

Něco

Odhlásit

Seznam děl

Filtr:

Jména autorů: Vše ▼

Full-text: Zadejte text

Rok vydání

od: 1812 ▼

do: 2016 ▼

Rozděláno ☒

Hotovo ☒

Zkontrolováno ☐

Aplikovat filtr

Zobrazit 10 ▼ záznamů

Vyhledat:

Název	Autor	Rok	Status	Akce	Přílohy
Almanach českého studentstva	Bohumil Adámek	1869	rozděláno	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (25)</a>
Na zemi a na nebi	Eduard Albert et al.	1900	hotovo	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (65)</a>
Básně	Ladislav Arietto	1892	hotovo	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (52)</a>
Na zemi a na nebi	Eduard Albert	1900	rozděláno	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (16)</a>
Když slunce zapadá	Karel Babánek	1900	rozděláno	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (89)</a>
Kniha písní	Karel Babánek	1908	rozděláno	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (57)</a>
Zrcadlení	Alfons Breska	1938	rozděláno	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (64)</a>
Ruce	Otokar Březina	1901	hotovo	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (24)</a>
Z hlubin věků	Xaver Dvořák	1905	hotovo	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (92)</a>
Kopřivy	František Hajniš	1853	hotovo	<a href="#">Metadata</a> <a href="#">Text</a> <a href="#">Smazat</a>	<a href="#">Přílohy (27)</a>

Zobrazeno 1 až 10 z 1700 výsledků

Předchozí

1

2

3

4

5

...

Další

Obrázek A.3: Hlavní stránka

35

## A. SCREENSHOTY

Seznam děl

Text

Přílohy

user@email.com ▼

### Údaje o dílu: Když slunce zapadá

#### Autor

Karel Babánek

Upravit

Odstranit

Karel Babánek

Upravit

Odstranit

Přidat nového

#### Ostatní

Titul:	Když slunce zapadá	Podtitul:	
Místo:	Praha	Vydavatel:	Weinfurter, Eduard; Stivín, Emanuel (Nákladem knihkupectví E. Weinfurtra v Praze. Tiskem E. Stivína v Praze.
Rok:	1900	Vydání:	1.
Stran:	32	Věnování:	
Moto:		Autor mota:	
Formát:		Zdroj signatura:	Národní knihovna ČR, Praha; 54 J 2058

Popis:

Exemplář je opatřen nepůvodní pevnou vazbou. Kartonové desky jsou potaženy papírem s žluto-hnědým mramorovaným vzorem, hřbet a rohy hnědým plátnem. V levém horním rohu předních desek je nalepen štítek s předtištěným rámem a signaturou vepsanou černým inkoustem. Hřbet nese svisle zlatenou ražbu 54 J 2058. Babánek, Když Slunce zapadá. a v hlavě a patě vodorovnou zlatou linku. Ořízku svazku zdobí ze tří stran modrý pavučinový vzor. Přední i zadní předsádku tvoří světlé zelený papír, v levém horním rohu předního předeštlí je černým inkoustem připsaná stávající signatura svazku. Kniha má zachovány přední stranu původní obálky z užího papíru s modrým potiskem, na vnější straně věvazanou před knižní blok. Zdobí ji secesní luneta, v níž sedí žena v říze a s vavřínovým věncem na hlavě; v horní části je obdélníkový rám s titulem sbírky, nad ním vlevo autorovo jméno, v patě menší luneta s údajem o roce ...

Ediční poznámka:

V diplomatickém zápisu Nakladatelských údajů přejímáme vyčerpávající znění tiráže ze strany [35]. Jazyk Babánkovy třetí sbírky vychází z jednoduché píšňové formy, ojedinělé odchylky od současného úzu jsou motivovány spíše dobovou normou než autorskou invencí. Ve shodě s našimi edičními zásadami text maximálně šetříme, abychom jeho zvláštnosti nesetřeli. Nezachováváme ovšem sazbu kurzivou, která má charakter čistě grafický, nikoli významový. Znak uvozovek »« nahrazujeme znakem běžných uvozovek. V oblasti hláskosloví dodržujeme kvantitu samohlásek (<i>víl</i>,<i>dvěfe</i>,<i>tony</i>,<i>altan</i>), hláskové výpustky (<i>cnosti</i>,<i>říc</i>) i nevokalizované předpony (<i>rozzpívalo</i>,<i>rozsmutnělé</i>). Ponecháváme rovněž nenoremní jevy pravopisné: velká písmena ve slovech, jež mají symbolický význam nebo jsou zvýrazněna (<i>Radost</i>), velké písmeno jako prostředek vyjádření úcty u slovesa<i>Jsi</i>, dále distribuci ...

Obrázek A.4: Metadata

Seznam dělTextMetadata

Přílohy: Na zemi a na nebi

Pořadí: 68 ▾  
Soubor nevybrán Vybrat soubor Nahrát

Pořadí: 1 ▾  
Změnit Smazat

Pořadí: 4 ▾  
Změnit Smazat

Pořadí: 3 ▾  
Změnit Smazat

Pořadí: 5 ▾  
Změnit Smazat

Pořadí: 5 ▾  
Změnit Smazat

Pořadí: 6 ▾  
Změnit Smazat

Předchozí 1 2 3 4 5 ... Další

user@email.com ▲

Změna hesla

Něco

Něco

Odhlásit

Obrázek A.5: Přílohy

## A. SCREENSHOTY

Seznam děl

user@email.com ▼

### Seznam autorů a vydavatelů

Zobrazit

10 ▼

záznamů

Vyhledat:

Jméno ▲	Příjmení ◆	Korporace ◆	Akce	
Acta universitatis Palackian...	neuvedeno	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
Alois	Lapáček	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
Augustin Eugen	Mužik	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
Ferdinand	Tomek	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
František	Šimáček	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
Knihovna Společnosti Petr...	neuvedeno	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
Jaroslav	Pospíšil	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
Josef	Pech	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
Gustav	Pfleger Moravský	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>
Moravskolezské tiskárny	neuvedeno	Tiskárny a.s.	<a href="#">Upravit</a>	<a href="#">Smazat</a>

Zobrazeno 1 až 10 z 1700 výsledků

[Předchozí](#)   1   2   **3**   4   5   ...   [Další](#)

Obrázek A.6: Seznam autorů a vydavatelů

# Text k dílu: Když slunce zapadá

tučné

kurzíva

strofa

strana

řádek

uvozovky

<titulsbirka><b>HYMNY A VZDECHY.</b></titulsbirka>
<br />
<strana id="1" netistena="ano">[1]</strana>
<sbirka id="1">
<nadpis><b>HYMNY A VZDECHY.</b></nadpis>
<br />
<podtitulsbirky><b>CYKLUS BÁSNÍ</b></podtitulsbirky>
<br />
<spisovatel><b>AUG. EUG. MUŽÍKA.</b></spisovatel>
<br />
<tiraz>V PRAZE.</tiraz>
<tiraz>KNIHTISKÁRNA F. ŠIMÁČEK, NAKLADATELÉ.</tiraz>
<tiraz>1892.</tiraz>
<br />
<strana id="2" netistena="ano">[3]</strana>
<basen id="1">
<nadpis><b>RANNĚMU SKŘIVANU.</b></nadpis>
<br />
<strofa id="1">
<v id="1">Jsi oživlý ty kvítek, záhy</v>
<v id="2">jenž z jamí vyšel chladné víahy</v>
<v id="3">v ty opuštěné, smutné dráhy</v>
<v id="4">a pustý dol,</v>
<v id="5">kde vzduch je dosud sotva vlahý</v>
<v id="6">a mrtvo kol?</v>
</strofa>
.
.
.
<strofa id="3">
<v id="1">Ó chtěl bych písni být, jež hučí českou zemí</v>
<v id="2">v den Krista vzkříšení, když kraj prost zimy pout,</v>
<v id="3">ó chtěl bych vítězně se vznésti nad bor lemy</v>
<v id="4">a zahynout.</v>
</strofa>
<strofa id="4">
<v id="1">Ó chtěl bych ležeti tu, žertva bojů slavných,</v>
<v id="2">hrud' křížem protatu a v srdci rány pal,</v>
<v id="3">ó chtěl bych mrtev být za návrat plesů dávných</v>
<v id="4">a cítit dál.</v>
</strofa>
<sifra></sifra>
<strana id="128" netistena="ano">[131]</strana>
</basen>
</sbirka>

Obrázek A.7: Text díla

39



## Seznam použitých zkratk

**GUI** Graphical user interface

**XML** Extensible markup language

**TEI** Text Encoding Initiative

**HTML** HyperText Markup Language

**JRE** Java Runtime Environment

**UČL AV** Ústavu české literatury Akademie věd České republiky

**ISS** Internet Information Services

**DRM** Digital rights management





## Obsah přiloženého CD

	readme.txt.....	stručný popis obsahu CD
	exe .....	adresář se spustitelnou formou implementace
	src	
	impl.....	zdrojové kódy implementace
	thesis .....	zdrojová forma práce ve formátu L <sup>A</sup> T <sub>E</sub> X
	text .....	text práce
	thesis.pdf .....	text práce ve formátu PDF
	thesis.ps .....	text práce ve formátu PS