

Software pro tvorbu e-knih

Jakub Hubáček

Bakalářská práce
2014



Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Univerzita Tomáše Bati ve Zlíně

Fakulta aplikované informatiky

akademický rok: 2013/2014

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Jakub Hubáček**
Osobní číslo: **A11107**
Studijní program: **B3902 Inženýrská informatika**
Studijní obor: **Informační a řídicí technologie**
Forma studia: **prezenční**

Téma práce: **Software pro tvorbu e-knih**

Zásady pro vypracování:

1. Vypracujte přehled základních principů elektronických knih, popište některé standardně užívané formáty e-knih.
2. Rámcově analyzujte současný stav řešené problematiky a existující software.
3. Navrhněte strukturu software pro stylizování a převod prostého textu do vybraných formátů e-knih.
4. Prakticky realizujte software podle návrhu z předcházejícího bodu.
5. Otestujte a odladte tento program na praktických příkladech.
6. Vypracujte jednoduchou uživatelskou příručku pro jeho použití.
7. Výsledný software publikujte na Internetu.
8. Navrhněte směry budoucího vývoje výsledného díla.

Rozsah bakalářské práce:

Rozsah příloh:

Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam odborné literatury:

1. **KINDLE DIRECT PUBLISHING. Building Your Book for Kindle.** Seattle: Amazon Digital Services, Inc., 2012. ISBN ASIN B007URVZJ6.
2. **DARWIN, Ian F. Java: kuchařka programátora.** Vyd. 1. Brno: Computer Press, 2006, 798 s. ISBN 80-251-0944-5.
3. **Amazon Kindle Publishing Guidelines [online].** Amazon.com, Inc., 2013 [cit. 2014-01-30]. Dostupné z: <http://kindlegen.s3.amazonaws.com/AmazonKindlePublishingGuidelines.pdf>
4. **BB eBooks [online].** 2013 [cit. 2014-02-05]. Dostupné z: <http://bbebooksthailand.com/index.html>
5. **KISZKA, Bogdan. 1001 tipů a triků pro programování v jazyce Java.** Vyd. 1. Brno: Computer Press, 2012, 530 s. ISBN 978-80-251-2467-3.

Vedoucí bakalářské práce:

Ing. Bc. Pavel Vařacha, Ph.D.

Ústav informatiky a umělé inteligence

Datum zadání bakalářské práce:

28. února 2014

Termín odevzdání bakalářské práce:

13. června 2014

Ve Zlíně dne 28. února 2014

prof. Ing. Vladimír Vašek, CSc.
děkan



prof. Ing. Vladimír Vašek, CSc.
ředitel ústavu

Prohlašuji, že

- beru na vědomí, že odevzdáním diplomové/bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomová/bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové/bakalářské práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji diplomovou/bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou/bakalářskou práci nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- beru na vědomí, že pokud bylo k vypracování diplomové/bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové/bakalářské práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové/bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- Že odevzdaná verze diplomové/bakalářské práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně

.....
podpis diplomanta

ABSTRAKT

Cílem této bakalářské práce je navrhnout a vytvořit program schopný převodu prostého neformátového textu do formátu elektronických knih. Teoretická část práce popisuje elektronické knihy a jejich principy. Na výsledcích trhu je ukázáno, jak se trh s elektronickými knihami vyvíjí v ČR. Praktická část práce se zabývá návrhem a naprogramováním programu schopného převodu textu do elektronické knihy. Výhodou programu je, že uživatel nepotřebuje znalosti HTML nebo XML k vytvoření e-knihy. Součástí je také uživatelská příručka k programu.

Klíčová slova: Java, HTML, XML, tvorba e-knih, KindleGen, e-kniha

ABSTRACT

The object of this bachelor thesis is project and construct software capable of converting plain unformatted text to format of electronic books. Theoretical part describes electronic books and their principles. Based on market result is show, how market with electronic books is evolving. Practical part deals with project and programming of software, capable of making electronic book from plain text. Benefit of software is unnecessary knowledge of HTML or XML to making e-book. Part of the work is also user's guide to the program.

Keywords: Java, HTML, XML, making of e-book, KindleGen, e-book

„Cenu má pouze cesta. Pouze ona trvá...“

Antoine de Saint-Exupéry

Děkuji vedoucímu bakalářské práce panu Ing. Bc. Pavlovi Vařachovi, Ph.D. za velmi cenné rady, připomínky a užitečnou metodickou pomoc, kterou mi poskytl při vypracování mé bakalářské práce.

Také bych chtěl poděkovat své rodině a přátelům za morální podporu při zpracování bakalářské práce.

OBSAH

| | |
|--|-----------|
| ÚVOD..... | 8 |
| I TEORETICKÁ ČÁST..... | 9 |
| 1 CO JE TO E-KNIHA | 10 |
| 1.1 HISTORIE E-KNIH | 11 |
| 1.2 TECHNICKÁ STRÁNKA ČTEČEK E-KNIH | 13 |
| 1.2.1 DRM..... | 13 |
| 1.2.2 E-ink | 14 |
| 1.2.3 OCR..... | 14 |
| 1.2.4 Struktura e-knihy | 15 |
| 1.3 ROZDÍLY VE FORMÁTECH E-KNIH | 15 |
| 1.3.1 EPUB..... | 15 |
| 1.3.2 MOBI | 16 |
| 1.3.3 AZW..... | 16 |
| 1.3.4 LIT..... | 16 |
| 1.3.5 PDF | 17 |
| 1.4 SITUACE E-KNIH V ČR..... | 17 |
| 2 ANALÝZA SOUČASNÝCH NÁSTROJŮ | 21 |
| 2.1 MICROSOFT WORD A KINDLEGEN | 21 |
| 2.2 LIBREOFFICE A WRITER2EPUB | 22 |
| 2.3 MOBIPOCKET..... | 22 |
| 2.4 SYGIL | 23 |
| II PRAKTICKÁ ČÁST | 24 |
| 3 NÁVRH PROGRAMU | 25 |
| 3.1 HLAVNÍ OKNO | 26 |
| 3.2 SPRÁVA A APLIKACE STYLŮ | 26 |
| 3.3 SPRÁVA APLIKOVANÝCH STYLŮ | 26 |
| 3.4 ZPROSTŘEDKOVÁNÍ PŘEVODU DO FORMÁTU E-KNIHY | 27 |
| 4 REALIZACE PROGRAMU | 28 |
| 4.1 MAINWINDOW | 28 |
| 4.2 JDIALOG POMOCNÁ OKNA | 29 |
| 4.2.1 Písmo..... | 29 |
| 4.2.2 Informace o knize..... | 30 |
| 4.2.3 Metadata..... | 31 |
| 4.2.4 Nastavení programu | 31 |
| 4.3 STYLEMANAGER..... | 31 |
| 4.4 TAGMANAGER | 32 |
| 4.5 ZPRACOVÁNÍ XML SOUBORU | 32 |
| 4.6 CONVERTOR | 36 |
| 4.6.1 XML Builder | 37 |
| 4.6.1.1 Popis OPF souboru | 37 |
| 4.6.1.2 Popis NCX souboru | 38 |
| 4.6.2 Převod do e-knihy | 39 |

| | | |
|----------|--|-----------|
| 5 | UŽIVATELSKÁ PŘÍRUČKA S PRAKTICKÝMI UKÁZKAMI | 40 |
| 5.1 | PRVNÍ KROKY | 40 |
| 5.2 | MENU | 41 |
| 5.2.1 | Informace o knize | 41 |
| 5.2.2 | Nastavení programu | 42 |
| 5.2.3 | Nastavení písma | 42 |
| 5.3 | PŘEVOD KNIHY | 42 |
| 6 | PUBLIKACE PRÁCE A JEJÍ BUDOUCNOST | 43 |
| 6.1 | DALŠÍ VÝVOJ | 43 |
| | ZÁVĚR | 45 |
| | SEZNAM POUŽITÉ LITERATURY..... | 46 |
| | SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK..... | 49 |
| | SEZNAM OBRÁZKŮ | 51 |
| | SEZNAM GRAFŮ | 52 |
| | SEZNAM PŘÍLOH..... | 53 |

ÚVOD

Cílem mé bakalářské práce je vypracování návrhu a realizace aplikace schopné z prostého neformátovaného textu, vytvořit elektronickou knihu na základě uživatelem aplikovaných stylů. Aplikace bude pro uživatele přehledná a jednoduchá na ovládání. Výstupem aplikace bude elektronická kniha ve standardu moderních e-knih.

V teoretické části se zaměřím na obecné principy a využití elektronických knih. Nastíním vývoj elektronických knih a zařízení pro jejich četbu. Porovnáím základní rozdíly formátů pro sdílení elektronických knih a zmíním jejich hlavní rysy. Zaměřím se konkrétně na český trh s elektronickými knihami a jeho vývoj za poslední roky. V závěru teoretické části analyzuji dostupné řešení dané problematiky tvorby e-knih a jejich nedostatky.

V praktické části zahrnu teoretický návrh programu a jeho realizaci. Teoretický návrh bude obsahovat rozlišení aplikačních vrstev a jejich samostatnost. Program realizuji na základě teoretického návrhu s ohledem na možné změny. Realizovaný program odladím pro bezproblémové použití u koncových uživatelů.

K programu poskytnu uživatelskou příručku, která bude základním zdrojem informací pro orientaci v programu. Uživatelská příručka poskytne uživateli informace o rozložení a ovládání aplikace na praktických ukázkách.

Realizovaný program, včetně uživatelské příručky publikuji na diskusních fórech komunit zabývajících se elektronickými knihami.

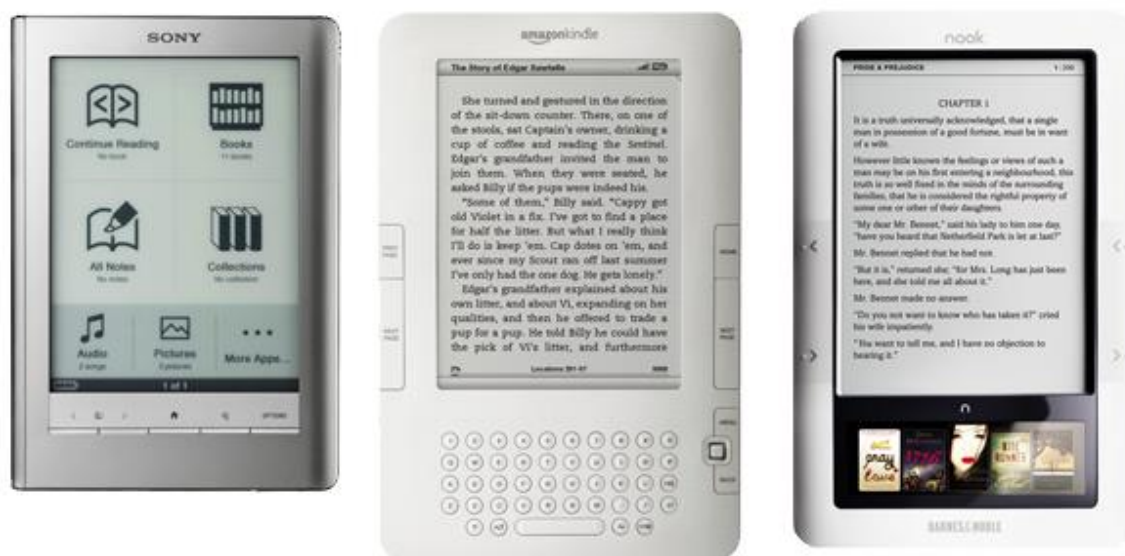
I. TEORETICKÁ ČÁST

1 CO JE TO E-KNIHA

E-kniha nebo zavedenější pojem e-book představuje digitální soubor s určitou strukturou, formátováním a obsahem. Pojem elektronická kniha zahrnuje nejen standardní formáty pro čtečky elektronických knih jako mobi nebo epub, ale také formáty pdf, txt, které nepatří mezi specifické formáty pro čtečky e-knih. Jako elektronickou knihu můžeme v podstatě označit jakoukoliv knihu, jejíž obsah byl digitalizován.

Čtečky e-knih, často mylně označované jako e-booky (nebo e-knihy), jsou zařízení, které jsou určeny pro zobrazování specifických formátů e-knih, zejména typy mobi a epub. V závislosti na preferencích výrobce jsou schopny zobrazovat vybrané typy formátů elektronických knih. Nejprve byla snaha prodejců prosazovat svůj vlastní formát e-knih. Avšak s přibývajícím subjekty trhu se snažili čtečky elektronických knih zacílit na ty nejčastěji používané formáty.

Obrovskou výhodou elektronických knih je jejich velikost. V tomto případě digitální velikost. Na průměrnou čtečku lze uložit 2000 elektronických knih, což představuje skutečně solidní knihovnu, zabalenou v zařízení velikostí podobnému jedné knize. Váha průměrné čtečky e-knih se nijak výrazně neliší od knihy papírové, většinou jsou čtečky lehčí. Dokonce se objevili výpočty hledající rozdíl mezi čtečkou e-knih naplněnou knihami a čtečkou, která je prázdná (na základě elektronů držených v tranzistorech paměti).[1]



Obr. 1 Zleva Sony PRS-600, Amazon Kindle 2 a Barnes & Noble Nook [2]

Stále existují autoři knih, jejichž výstupní formát není elektronický, ale odevzdávají rukopis nakladateli (John Irving – Svět podle Garpa). To je spíše výjimkou, většina spisovatelů využívá počítače a výhody textových editorů, ať už kvůli rychlosti psaní nebo jednoduchosti oprav.

1.1 Historie e-knih

Snaha o digitalizaci knih začíná v momentě, kdy si lidstvo uvědomuje, že digitalizovaný obsah je snadnější procházet, zpracovávat a přenášet. Zpočátku se knihy přepisovaly, poté vznikl software, který dokázal z naskenovaného obrazu strany rozpoznat písmena. Nejjednodušší cesta je mít knihu přímo v elektronické verzi dříve než jde k nakladateli.

V roce 1946 se v hlavě italského jezuita Roberta Buja rodí myšlenka na Index Thomisticus. Tento nástroj má sloužit pro vyhledávání digitalizované práce Tomáše Akvinského. Projekt financoval Thomas Watson (zakladatel IBM). Po třiceti letech práce vychází Index Thomisticus v 56 svazcích. To už dávno neobsahuje jen práce Tomáše Akvinského, ale i spoustu autorů ze stejného období nebo s určitou vazbou na Tomáše Akvinského. Později vychází jak na CD, tak na DVD a dokonce má svoji webovou verzi.[3]

Velmi důležitý projekt, který dal směr elektronickým knihám, byl Projekt Gutenberg. Projekt Gutenberg si dal za cíl digitalizovat, archivovat a zpřístupňovat knihy ve výsledné elektronické podobě co nejširšímu okruhu lidí. Projekt založil Michael Stern Hart v roce 1971. Nejprve se snažil o digitalizaci literatury, na kterou vypršela autorská práva. Později sami autoři s nízkými prodeji začínají přispívat svými pracemi. Zpočátku šlo o prostý přepis knih, později se využívá programů na převod naskenovaného textu do elektronické podoby. V současné době nabízí Project Gutenberg 45 tisíc elektronických knih v různých formátech. [4]

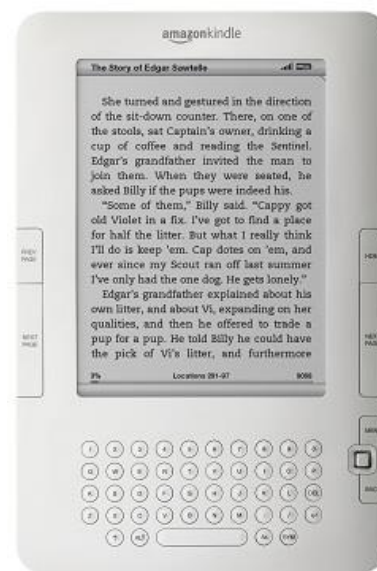
První čtečka e-knih v podobě jak je známá dnes, tedy zařízení, které padne do ruky, se objevila v roce 1998. V tomto roce byly vydány čtečky Rocket eBook firmou NuvoMedia a čtečka SoftBook firmou SoftBook Press, Inc. Obě zařízení mají dostupný svůj vlastní internetový obchod s elektronickými knihami. V případě čtečky SoftBook Press nebylo potřeba počítač pro stahování knih, stačilo do ní připojit RJ11 a interní modem čtečky e-knih se dokázal připojit k internetovému obchodu s knihami. V roce 2000 se firmy SoftBook Press a NuvoMedia spojují pod firmou Gemstar, která skoupila obě firmy a vydává Gemstar eBook. Gemstar eBook má také internetový obchod s knihami, listování, záložky a výdrž až 5 hodin, přesto v roce 2003 končí prodej jak čteček elektronických knih, tak knih z internetového obchodu. [6]



Obr. 2 Jedna z prvních čteček
- čtečka RCA [5]

V roce 2003 americké knihovny začaly s výpůjčkami e-knih. Literatura, kterou knihovny půjčovali, byla hlavně technického rázu. Nejvíce e-knih bylo z technických oborů, postupně se přidaly další obory a dokonce literatura fikce a beletrie.

Zlom v prodeji nastává v roce 2007, kdy Amazon vydává čtečku e-knih Kindle. Sice se odvětví s elektronickými knihami posunovalo dopředu, ale nikdo neměl pořádnou podporu čtenářů. Amazon, jehož hlavním artiklem byly právě knihy, má tedy nejlepší předpoklady pro inovaci, zkušenosti se čtenáři a také jeho obchod s knihami působí v různých částech světa.



Obr. 3 Amazon Kindle
2. generace [7]

Sony v roce 2009 zavádí půjčování knih z knihoven pomocí sítě Overdrive, která umožní výpůjčky elektronických knih. Sony se také snažilo zapojit čtečky e-knih více do škol.

V roce 2009 vévodí všem čtečkám elektronických knih Kindle od Amazonu, nejspíše díky dostupnosti a široké nabídce obchodu Amazon. Velké zahraniční nakladatelství se postupně připojují a začínají vydávat své knihy jak v papírově formě, tak ve formě elektronické knihy. S růstem trhu přichází

další hráči jako Asus, Samsung, Wooky nebo Prestigio. Amazon od roku 2011 prodává dokonce více e-knih než knih papírových. [8]

V roce 2011 se objevuje první česká čtečka elektronických knih. Vydává ji společnost eReading.cz a je postavena na čtečce Hanlin V60 jen s pozměněným firmware pro český trh. Čtečka eReading chce konkurovat hlavně Amazonu a může, protože Amazon neprodává české e-knihy. Tuto skulinu se snaží eReading uzavřít a propojit se svým obchodem eReading.cz dostupným přímo z čtečky eReading.cz. V dalších verzích své čtečky e-knih zpřístupňuje dokonce výpůjčky e-knih a k tomu potřebný firmware. Svou českou lokalizací a českým obchodem dostupným přímo z čtečky e-knih si získala velké sympatie. [9]



Obr. 4 Čtečka eReading.cz [10]

1.2 Technická stránka čteček e-knih

Čtečky e-knih se na rozdíl od tabletů dělají přesně pro potřeby e-knih, alespoň taková myšlenka původně byla. Například požadavky na displej čteček e-knih jsou úplně jiné. Displej s milionem barev se stává zbytečným, pokud potřebujeme přečíst jen text. Nedostatek barev čtenáři rádi nahradí zvýšenou dobou výdrže.

Rozšíření internetového pirátství na e-knihy donutilo vydavatele použít zabezpečení, které šíření zabrání. Ochrana DRM je na jednu stranu funkční, ale na druhou má své nedostatky.

1.2.1 DRM

„Zkratka DRM pochází z anglického Digital Rights Management - správa digitálních práv. DRM neboli také „tvrdé DRM“ je způsob jak kontrolovat nebo omezovat používání digitálního obsahu. Rozsah DRM ochrany většinou stanovuje nakladatel. U e-Knih představuje vždy větší nebo menší omezení, která limitují počet kopírování, tištění nebo sdílení. Často je omezeno na stažení pouze jedné kopie a to dokonce pouze jednoho formátu, který si při nákupu vyberete. Běžně se lze setkat také s modelem pouze jedné licence vázané k IP adrese.“[11]

DRM lze taky nastavit tak, aby obsah byl dostupný jen pro jeden konkrétní účet, vázaný na obchod s e-knihami. Všechny tyto omezení, ale uživatele velmi omezují a kvůli klesajícím prodejům e-knih s ochranou DRM se hledají jiné způsoby.

1.2.2 E-ink

E-ink je technologie, díky které vydrží čtečky knih daleko více hodin provozu, než kdyby měli displej na technologii LCD. E-ink displej spotřebovává energii jenom změnu zobrazení, to znamená přepnutí stránky. Kindle 4 vydrží údajně při čtení půl hodiny denně až měsíc na jedno nabití.

E-ink technologie je založená na přesunu mikročástic v průhledných mikro-kapslích. Mikročástice jsou dvojího druhu: bílé a černé, nabité opačným nábojem. Průhledná dutinka je naplněna bezbarvým médiem.

„Ke změně stavu jednoho bodu (jedné mikrokapsle) je potřeba el. proud o velikosti asi 20 nA, tj. necelý 1 mikroampér na centimetr čtvereční. Tedy zhruba jen tisícinu energie, než kolik jí potřebují monitory nebo displeje. K "tisku" dokonce není potřeba ani vlastní zdroj energie, protože stav mikročástic lze měnit i působením elektromagnetických vln.“ [12]

1.2.3 OCR

Převodu knihy z její papírové podoby do elektronické podoby lze dosáhnout dvěma způsoby. První způsob spočítá v tom, že máme digitalizovaný obsah k dispozici, tedy ho můžeme jen upravit pro čtečku e-knih a je hotovo. Pokud ovšem nemáme takový digitalizovaný obsah, například velmi staré knihy, musíme obsah získat pomocí OCR.

Optical Character Recognition neboli Optické rozpoznávání znaků je způsob, jak z naskenovaného obrazu textu knihy získat digitalizovaný text. Program porovnává znaky na vstupním obrazu se znaky, které jsou mu dostupné. Při porovnávání si program stanovuje pravděpodobnost znaku na vstupu se znakovou sadou. Poté vezme znak ze znakové sady s největší pravděpodobností a ten se objeví na výstupu.

Programy pro OCR jsou nyní už velmi vyspělé, rozeznávají i diakritiku nebo znaky z ruské abecedy. Kromě prostého rozpoznávání znaků se do programů integrují slovníky, které dodatečně kontrolují převedený text. Chyby se přesto objevují a nejčastěji dělená pomlčkou.[13]

1.2.4 Struktura e-knihy

Formáty souborů pro e-knihy nejsou jen souborem, ve kterém nalezneme stylování nebo text v jednom jediném souboru, jak to je například u HTML stránky s vnořeným CSS souborem.

Formáty e-knih jako EPUB a MOBI jsou rozdělené na 3 části:

- Část obsahující pouze neformátovaný text.
- Část obsahující vlastnosti jednotlivých textových úseků.
- Dodatečné informace o knize, jako je jméno autora nebo obrázky.

Základní problém, proč dokumenty jako DOC nebo PDF nemůžou být jen tak otevřeny v čtečce e-knih, je ten, že se text těchto souborů nedokáže jen tak přizpůsobit šířce čtečky elektronických knih. Některé čtečky e-knih kompatibilitu s formáty DOC a PDF mají. Většinou to znamená, že text si bude muset uživatel přiblížit. [14]

1.3 Rozdíly ve formátech e-knih

Projekt Gutenberg, který vytvářel největší podíl na přírůstu čteček e-knih, má většinu e-knih formátovaných do otevřených formátů jako je plain text, html nebo pdf. Nejdříve ze všeho se e-knihy ukládali do plain text formátu, který je nejjednodušším formátem pro uživatele. Formát HTML přibyl s rozšířením základny dobrovolníků pracujících na tomto projektu. S růstem dobrovolníků vzniká prostor pro úpravu knih do dalších formátů dokumentů, jako je například PDF.

Když na trh knih vstoupili čtečky elektronických knih a samotné elektronické knihy, vzniklo množství formátů na základě preferencí prodejců. Kindle preferuje Mobi, Nook od Barnes & Noble preferuje pdb, Wooky má formát Wooky atd. Mezi nejpopulárnější formáty e-knih podle mě patří zejména: EPUB, MOBI, AZW, LIT a PDF. I přesto, že PDF formát není přímo formát pro e-knihy, stal se velmi používaným a rozšířeným i na čtečkách e-knih.

1.3.1 EPUB

Otevřený formát vytvořený pro e-knihy organizací International Digital Publishing Forum. Jeho koncovka .epub lze interpretovat jako .zip a teprve struktura obsahu z něj dělá formát pro e-knihy. Obsah je zapsán pomocí XML a HTML, cílem bylo vytvořit, jak příjemně programovatelný formát, tak formát, který bude jednoduše dostupný pro koncového uživatele.

V českém prostředí se jedná o nejpoužívanější formát. Ochrana DRM zpočátku chyběla i ta byla přidána v pozdějších verzích.

Podporován na čtečkách: B&N Nook, Kobo, Sony, Wooky Reader, eReading.cz

1.3.2 MOBI

Od formátu EPUB se liší v podstatě jen větší mírou komprese textu. Vznikl původně jako součást programu Mobipocket Reader pro zařízení PDA. Přes zařízení PDA se rozšířil MOBI formát postupně na chytré telefony i na počítače. MOBI má ochranu pomocí DRM a dokáže spolupracovat se zabudovaným slovníkem. Společnost Mobipocket byla v roce 2005 koupena společností Amazon. Většina knih v českých i zahraničních obchodech MOBI většinou podporuje, dle mého je to kvůli velké rozšířenosti zařízení Kindle.

Struktura MOBI souboru:

- Seznam souborů (OPF soubor) definuje informace o knize, umístění všech souborů, kostru knihy a odkazy pro rychlý přístup k důležitým prvkům knihy.
- Seznam odkazů (NCX soubor) obsahuje odkazy na vybrané části dokumentu, většinou například kapitoly. Na základě odkazů pak dokáže odkazovat na kapitoly v knize.
- Ostatní soubory (IMG, HTML, ...) jako obsah, obálka. Sem patří všechny soubory čitelného nebo zobrazitelného obsahu e-knihy.

Podporován na čtečkách: B&N Nook, Kindle, eReading.cz

1.3.3 AZW

Formát speciálně pro Amazon Kindle. V podstatě dříve používaný formát MOBI a přidává do něj kompresi, tím vzniká formát AZW. Oproti staršímu MOBI formátu AZW má ochranu pomocí DRM. Tento formát je možné sehnat jen nákupem na Amazonu a přenosem do čtečky přes síť Whispernet (zprostředkovatel přenosu e-knihy do čtečky e-knih).

Podporován na čtečkách: Kindle

1.3.4 LIT

Formát čitelný pouze programem Microsoft Reader. Tento formát používá DRM ochranu. Nikdy se pořádně nepoužíval pro e-knihy, zůstal u zařízení Windows Mobile a PDA zařízení

PocketPC od Microsoftu. Podpora formátu LIT a programu Microsoft Reader byla ukončena v roce 2011, to vedlo k následnému konci prodeje e-knih s formátem LIT.

Podporován na: PocketPC (PDA)

1.3.5 PDF

Vznikl pod společností Adobe Systems původně pro použití v produktech Acrobat. Ukládání e-knih do tohoto formátu se stalo praktickým hlavně kvůli tomu, že jeho programy pro jeho čtení jsou velmi dostupné. Obrovskou nevýhodou je pevná šířka stránky PDF dokumentů, kvůli tomu se musí v čtečkách většinou přiblížit část stránky a poté je teprve možné číst. Formát PDF také není možné znovu otevřít a upravit běžným způsobem.

Struktura PDF:

- Hlava obsahuje pouze jediný řádek a to informaci o verzi PDF
- Tělo obsahuje informace o všech obsažených nepřímých objektech v dokumentu. Například obrázky, fonty, slova, záložky nebo formuláře a další.
- Tabulka křížových odkazů je plná ukazatelů na všechny použité nepřímé objekty v dokumentu. Pro každý jeden objekt uchovává v tabulce informaci o jeho pozici v rámci dokumentu.
- V patě dokumentu je uložena pozice, na které začíná tabulka odkazů. [15]

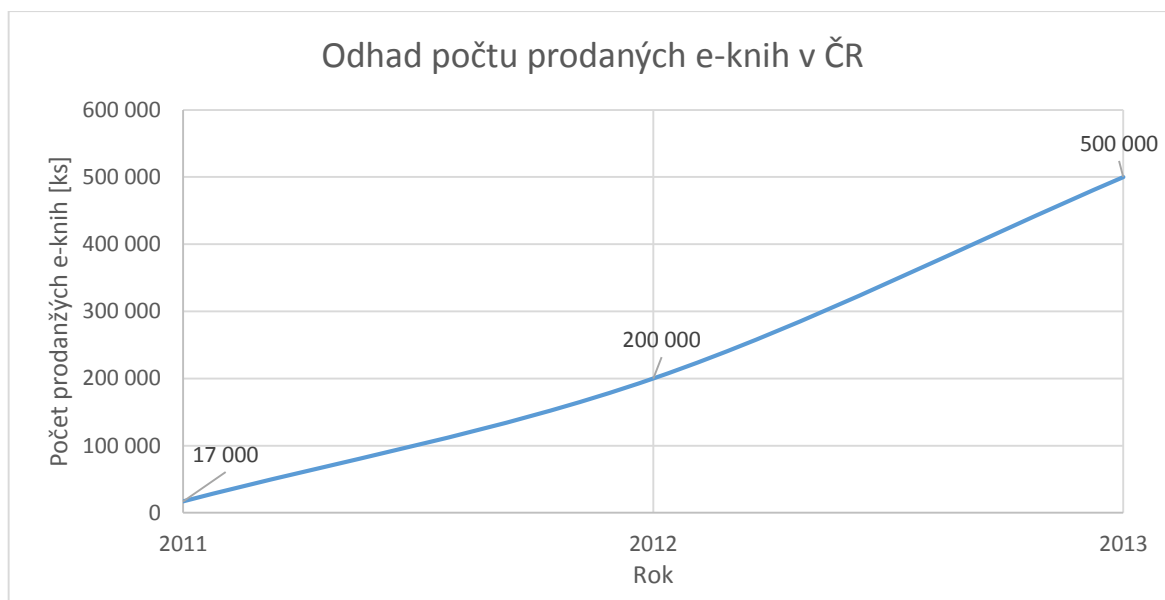
Podporován na čtečkách: B&N Nook, Kobo, Sony, Wooky Reader, eReading.cz [16,17,18]

1.4 Situace e-knih v ČR

V českém prostředí se jako první česká čtečka e-knih objevila eReading.cz, kterou brzy následovali čtečky Wooky a Palmknihy.cz. Tyto obchody reagovali na jasný nedostatek – přestože měl Amazon zavedený obchod a dokázal spolehlivě distribuovat knihy, neměl české knihy. České čtečky e-knih, tedy čtečky, které byly vydané pod některým ze zavedených českých prodejců e-knih, měli přímou cestu na obchod s elektronickými knihami. [16]

Trh s elektronickými knihami se v České republice stal během pár let plnohodnotným trhem. O tom svědčí i množství nových prodejců e-knih. Přestože se čtečky elektronických knih kvalitou většinou nevyrovnají zahraničním produktům, našli si svou cestu k zákazníkům díky dostupnosti českých knih. Z čtečky e-knih lze projít katalog internetového obchodu a taky zrovna elektronickou knihu zakoupit.

„Většina distributorů e-knih sází na otevřené formáty e-knih (tedy ePub a Mobi), které je možné číst prakticky na libovolném čtecím zařízení. Knihy jsou chráněny nejčastěji pomocí tzv. sociálního DRM, tedy vodoznakem vloženým do knihy a odkazujícím na konkrétního kupujícího. Výjimkou je platforma Wooky, která používá vlastního šifrovaného formátu dat pro vlastní čtečky.“ [19]

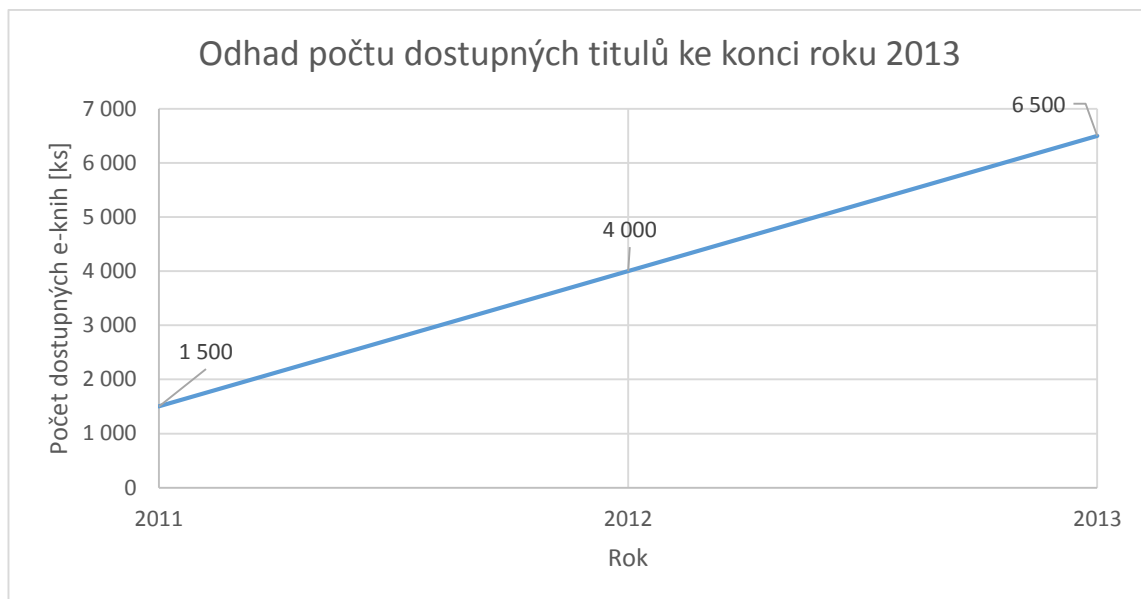


Graf 1 Odhad počtu prodaných e-knih v ČR [19]

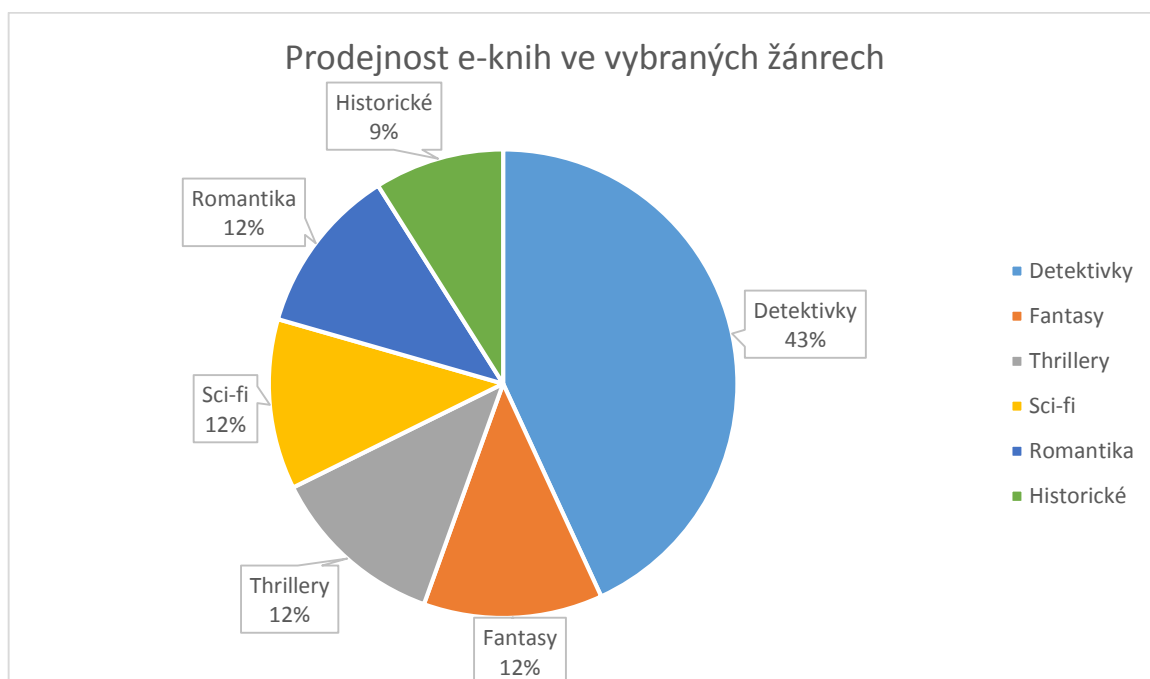
Jde vidět, že nárůst počtu prodaných e-knih je obrovský. Růst je stále velmi vysoký, i když se neočekává, že takový šikmý růst vydrží dlouho. Rok 2012 přinesl v porovnání s rokem 2011 podle odhadů až desetinásobný rozdíl prodeje elektronických knih.

„Celkově v roce 2012 vyšlo ve formě e-knihy zhruba 15 % vydaných papírových knih.“ [19]

V nejbližších letech se přesto neočekává, že by papírové knihy mohutně nahradili elektronické čtečky. Producenti e-knih vydávají obrovské množství e-knih. Někteří autoři začínají a zůstávají přímo u e-knih, ať už kvůli nákladům nebo šancím na mladém trhu.



Graf 2 Odhad počtu dostupných titulů ke konci roku [19]



Graf 3 Prodejnost e-knih ve vybraných žánrech [19]

Průměrná cena elektronické knihy se pohybuje okolo 145 Kč s DPH. Mezi největší nakladatele e-knih patří Kniha Zlín, Euromedia Group, Argo nebo Domino. Celkový počet producentů se však blíží ke dvou stovkám a více. Spousta lidí se také pouští do publikace e-

knihy na vlastní pěst, kdy si zařídí jak převod do formátů pro čtečky e-knih tak o prezentaci a prodej e-knihy. [19]

V posledním roce začal elektronické knihy nabízet i Google na svém internetovém obchodu Google Play. Nejdříve byla nabídka pouze zahraničních děl, ale v poslední době se objevují i české knihy. Google experimentoval s ochranou knih, která má posloupnost obsahu svázanou s Google účtem. Tedy při předání e-knihy na jiný účet by byla e-kniha nečitelná – přeházela by svůj obsah.

Půjčování e-knih v ČR v současnosti nedosahuje takové šíře nabídky jak v zahraničí. Největším problémem je, že nakladatelům takto odchází zisky z prodeje e-knih. Pokud tedy autor a nakladatel souhlasí s výpůjčkou e-knih zavádí se většinou něco, jako poplatek za období, ve kterém může být e-kniha prodejcem distribuována k půjčce. Z takové e-knihy nedostává nakladatel zisk, pouze poplatek na celé období.

V případě knihoven je nechuť nakladatelů ještě větší. Knihovny historicky půjčovali knihy a za danou výpůjčku neodváděli poplatky. V případě e-knih není, ale knihovna limitována fyzickými kusy k půjčení, tedy může půjčit tisíce e-knih a nakladatel teoreticky přijde o tisíce prodaných knih. To se nakladatelům samozřejmě nelíbí a v poslední době vzniká model paušálu – na určité období je možné půjčit určitý počet e-knih za určitou cenu. [20]

Celková cena e-knihy oproti papírové knize se zmenší o polovinu, zvýší se jen sazba DPH u e-knihy (považována za software). Cenu papírové knihy tvoří odhadem tyto položky:

- 30 % pro knihkupce (skladování a prezentace papírové knihy)
- 10 % pro distributora (nabídka knih knihkupcům, propagace)
- 30 % náklady na tiskárnu (tisk, vazba)
- 8 % pro autora (autorský honorář)
- 10 % tvoří DPH (u papírových je 15% sazba DPH, u e-knih je 21% sazba DPH)
- 12 % nakladatel (obálka, korektura, popřípadě překlad)

U elektronické knihy odpadá provize knihkupce a zanikají náklady na tisk knihy. E-kniha ve výsledku teoreticky stojí o 60% méně než kniha papírová. Bestsellery a knihy populárních autorů budou mít samozřejmě cenu závislou také na velikosti poptávky trhu. [21]

2 ANALÝZA SOUČASTNÝCH NÁSTROJŮ

V současné době se nejčastěji používá více kroků pro dosažení upravené elektronické knihy na výstupu. Nejvíc přímý způsob je nasylovat si HTML soubor podle sebe, přidat mu obálku a další informace o knize jako jsou autor nebo název knihy. Takto připravený soubor se pak pomocí generátoru jako KindleGen převede do finálního MOBI nebo EPUB souboru. [22]

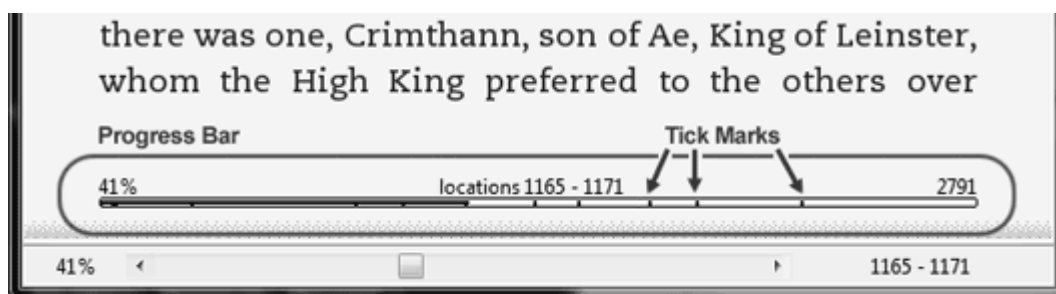
Od těchto způsobů se upouští, zejména kvůli dostupnosti nástrojů, které jsou přehledné a příjemné na práci. Nejpoužívanější jsou nyní nástroje, ve kterých nastyluje uživatel text a nástroj takový text převede do HTML. Cílem je, aby HTML bylo, co nejčistší. Výstupní HTML se vyčistí a dodatečně upraví. Na základě vyčištěného HTML se vytvoří obsah dokumentu, seznam souborů a seznam odkazů na kapitoly.

2.1 Microsoft Word a KindleGen

Microsoft Word patří mezi nejpoužívanější textové procesory. Patří do kancelářského balíčku Microsoft Office. Dokáže vytvářet a upravovat široké množství dokumentů, například DOC, DOCX, ODT nebo XML dokumenty. Nepoužívá se pouze pro tvorbu a editaci souborů, čistě založených na textovém obsahu, ale také na tvorbu vizitek, pohlednic, koláží nebo firemních letáků.

Kromě toho dokáže převést dokument do HTML formátu s uchováním aplikovaných stylů a rozložení. Výstup HTML se může jevit lehce nepořádný, ale pro uživatele bez znalosti HTML je to ideální cesta jak získat HTML formát. [23]

Získaný HTML formát je pak možné převést do elektronické knihy. Takový formát neobsahuje navigační body a přidání informací o knize nelze provést prostřednictvím aplikace Microsoft Word.



Obr. 5 Navigační body (Tick Marks) v čtečce Kindle [24]

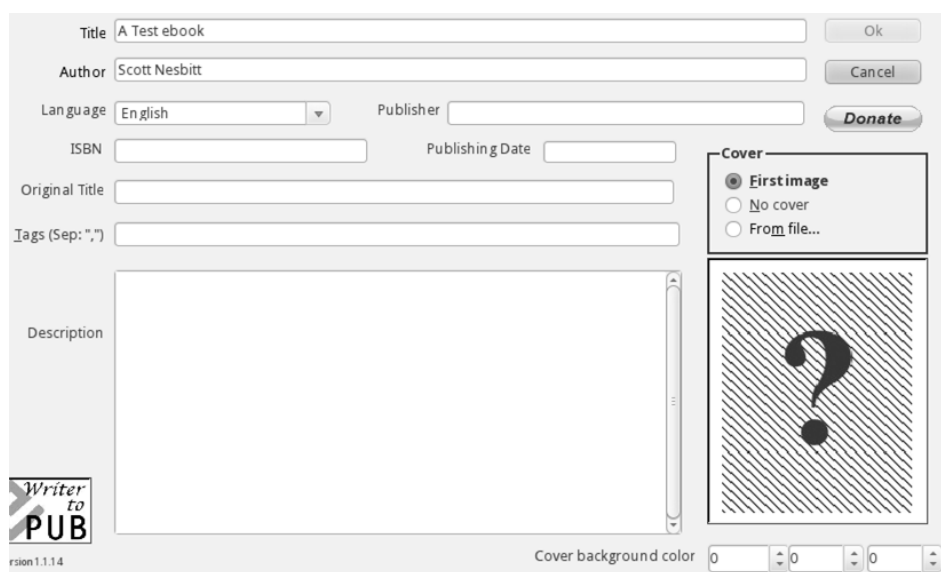
2.2 LibreOffice a Writer2epub

„LibreOffice je svobodný kancelářský balík pro Windows, Macintosh a Linux, který nabízí šest plnohodnotných aplikací pro práci s vašimi dokumenty a daty: Writer, Calc, Impress, Draw, Base a Math.“ [25]

LibreOffice patří mezi kancelářské balíčky, konkurujícím například kancelářskému balíčku od firmy Microsoft, má velmi podobné funkce pro úpravy textu nebo tabulek. V současné době je pro Windows dostupná verze 4.2.4, obsahující i českou lokalizaci.

LibreOffice Writer dokáže podobně jako Microsoft Word uložit vytvořený dokument do HTML. Při stažení doplňku nazvaného Writer2ePub může uživatel přímo v programu Writer nastavit informace o knize a vybrat obálku knihy.

Stále nepodporuje navigační body v elektronické knize, ale jeho doplněk Writer2ePub dokáže přidat základní informace o knize. [26]



Obr. 6 Okno s nastavením informací v programu Writer [26]

2.3 Mobipocket

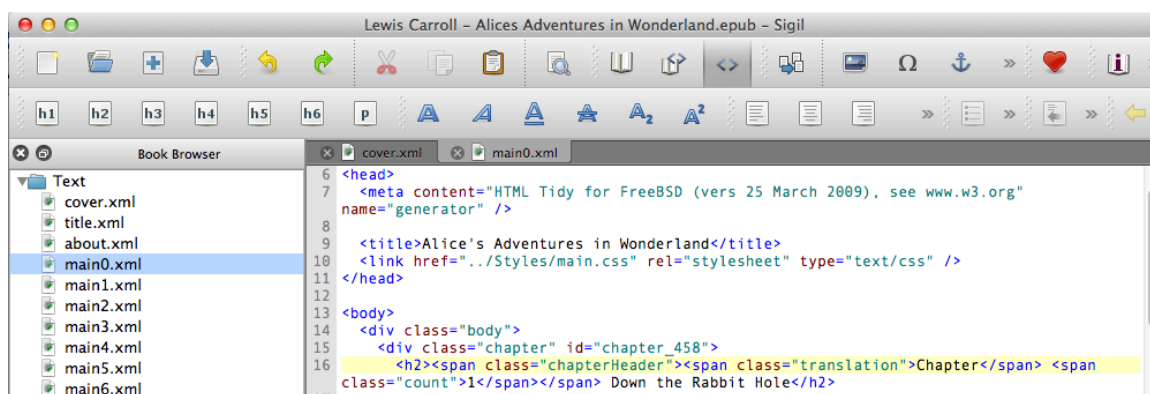
Mobipocket je původem francouzská společnost, která vytvářela a distribuovala knihy pro PDA zařízení. Pro účel distribuce elektronických knih byl vydán program Mobipocket Reader, schopný číst MOBI i PDB soubory.

Mobipocket je nyní ve vlastnictví společnosti Amazon. Zejména kvůli formátu e-knih MOBI. Program Mobipocket Creator je schopný na základě vloženého dokumentu DOC nebo TXT, vygenerovat elektronickou knihu ve formátu MOBI.

Nechybí přidání informací o knize nebo nastavení obálky knihy. Stránka s obsahem se tvoří na základě stylů z dokumentu DOC. Převod do formátu EPUB nepodporuje. [27]

2.4 Sygil

Sigil je velmi pokročilý nástroj pro vytváření e-knih ve formátu EPUB. Dokáže spravovat stránku s obsahem i metadata a obálku knihy. Uživatelé nabízí spoustu nástrojů pro tvorbu e-knihy. Právě díky jeho otevřeným možnostem je vhodný pro uživatele se znalostí struktury e-knih a formátů jako XML, OPF a NCX. Výsledné knihy převádí do EPUB, MOBI nepodporuje. Sigil mě příjemně překvapil a doporučuji ho vyzkoušet.



Obr. 7 Editace knihy v aplikaci Sigil [28]

Většina nástrojů tedy funguje na převodu HTML do e-knihy. Ať už programy převádí pouze z HTML formátu nebo používají i další formáty rozšiřující knihu o informace, vždy potřebujeme několika kroků.

Mým cílem je vytvořit aplikaci, u které nebude uživatel potřebovat znalosti formátů HTML, XML nebo CSS. Aplikace by se měla ve výsledku postarat o převod jako i o nastavení informací o knize.

II. PRAKTICKÁ ČÁST

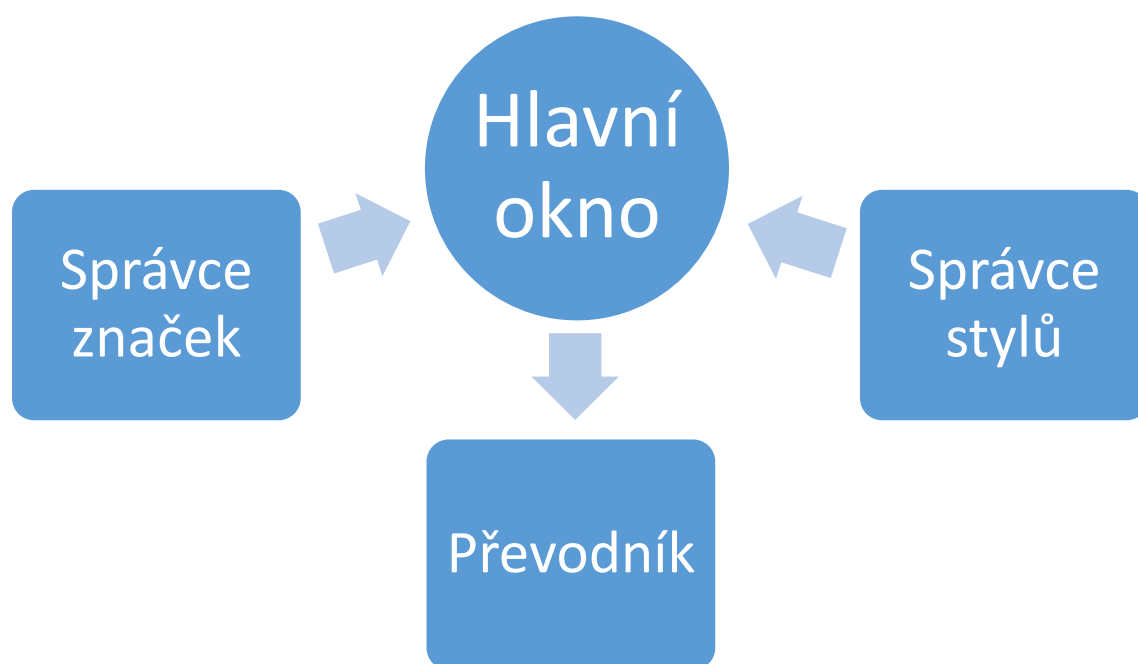
3 NÁVRH PROGRAMU

Program bude určený pro uživatele, kteří chtějí vytvořit e-knihu a jsou jen průměrně zdatní v práci na PC. Není potřeba znalosti HTML pro obalování textu pomocí tagů, ani není potřeba zkušenosti s CSS pro stylování jednotlivých částí dokumentu.

Každá třída bude mít svůj vlastní význam. Třídy budou oddělené tak, aby žádná třída nezasahovala, co se obsahu týče, do jiné třídy. Hlavní okno by mělo tedy jen reprezentovat data. Celá úprava a zpracování dat bude probíhat v přidružených třídách. Výsledky volaných metod v přidružených třídách budou předávány zpět do hlavního okna. Uživatel tak bude informován, jestli volaná akce skončila úspěchem či nikoli.

Obecný návrh tříd:

- Hlavní okno
 - Spravuje všechny události a předává data kam má.
- Správce značek
 - Spravuje značky, které odpovídají aplikovaným stylům.
- Správce stylů
 - Aplikuje vybrané styly na vybraný dokument.
- Převodník
 - Na základě nastavených dat a aplikovaných stylů vytvoří e-knihu.



Obr. 8 Vizualizace práce programu [vlastní]

3.1 Hlavní okno

Hlavní požadavek na hlavní okno aplikace bude jeho jednoduchost a ovladatelnost. Čím více „user-friendly“ tím lépe pro koncového uživatele. V hlavním okně bude možnost načíst text reprezentující e-knihu a ten aplikovat různé styly.

Ovládací prvky pro aplikaci stylů budou rozmístěny tak, aby pro uživatele nebylo obtížné najít požadované styly a ty následně aplikovat. Místo klasických stylů, jako tučné písmo nebo kurzíva, bude taky možné na text aplikovat i styly pro různé typy nadpisů nebo zvýraznění.

Kromě vizualizace aplikovaných stylů bude možné přepnout zobrazení, kde uživatel uvidí přímo text obalený pomocí párových značek.

3.2 Správa a aplikace stylů

V této části budu aplikovat styly textu aplikace, tak aby byli vhodně vizualizovány. Styly by měli být nastavené tak, jak je na ně uživatel z běžných prostředí zvyklý. Hlavní činností této části aplikace bude tedy přiřazování stylů na text.

Aplikace bude od hlavního okna dostávat 2 typy stylů – jednoduchý a komplexní.

- Jednoduchý stylem chápu styl jako: Tučné, Kurzíva, Podtržení. Takový styl bude vždy nastavovat jen jednu vlastnost – velikost, písmo nebo tučnost písma.
- Komplexní styl bude představovat například nadpis určité části dokumentu nebo celý odstavec. Odstavec, na který použiji komplexní styl, bude například kurzívou a zároveň tučným fontem.

3.3 Správa aplikovaných stylů

Při aplikování stylu se tato část aplikace postará o zaznamenání daného stylu. Bude si vést záznamy ve všech aktuálních aplikovaných stylech. Cílem bude, aby si uživatel mohl zobrazit, jaké styly na konkrétní části textu použil.

Z hlavní aplikace bude vždy zaslán požadavek na konkrétní změnu na konkrétním rozsahu textu. Rozsah a požadovaný styl se předá do aplikace zodpovědné za změnu stylů. Aplikace zodpovědná za změnu stylů následně vyvolá požadovanou změnu dokumentu.

3.4 Zprostředkování převodu do formátu e-knihy

Finální část aplikace bude poskytovat převod do vybraného formátu elektronické knihy. Uživatel zvolí místo pro uložení elektronické knihy a aplikace se postará o překlad. Obsah pro překlad do elektronické knihy bude brát z části aplikace, kde jsou styly zaznamenány.

Při překladu se kromě běžných prvků, jako odstavce a nadpisy, objeví i prvky reprezentující konec stránky elektronické knihy nebo další informace o knize.

Informace o knize, jako její autor nebo název knihy, bude možné nastavit. Nesmí chybět zvolení obálky elektronické knihy. Kromě informací o knize a volbě obálky knihy bude moci uživatel vybrat také, zda chce generovat seznam kapitol v knize. V seznamu kapitol budou zahrnuty všechny kapitoly spojené hypertextovými odkazy mezi HTML dokumenty.

Při formátování obsahu k převodu se budu řídit doporučeními danými k formátování e-knih. HTML připravené pro převod bude obsahovat kromě známých CSS a HTML elementů také CSS vlastnosti využívané zejména elektronickými knihami. [22]

4 REALIZACE PROGRAMU

Hlavní třída programu, která bude mít na starosti všechny ostatní třídy, jejich metody, volání a výstupy, bude MainWindow. Přes třídu MainWindow probíhá komunikace mezi všemi ostatními třídami – předávají se nastavené hodnoty, informace o knize nebo lokaci k uložení vytvořené e-knihy.

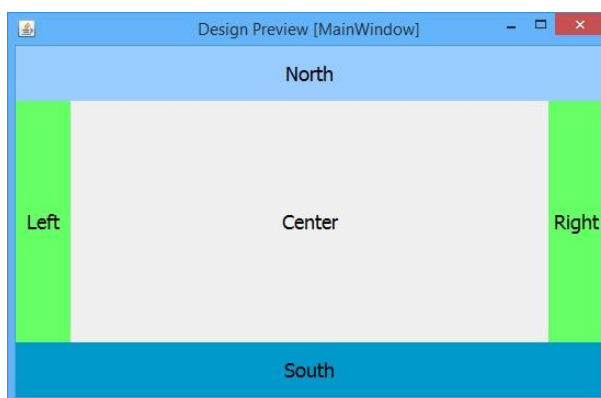
Okolo hlavního okna existuje mnoho dalších částí obslužných oken, které budou sloužit pro nastavení dílčích částí programu.

4.1 MainWindow

Rozvržení aplikace jsem vybíral tak, aby rozvržení jednotlivých prvků bylo zejména přehledné a dalo se s ním dobře manipulovat. Pro rozvržení ovládacích prvků jsem vybral BorderLayout.

BorderLayout je rozvržení, které rozdělí aplikaci za 5 částí:

- severní část (může být vertikálně zvětšována směrem dolů)
- jižní část (může být vertikálně zvětšována směrem nahoru)
- západní část (může být posunuta pravou hranicí horizontálně)
- střední část (může být posunuta jak vertikálně, tak horizontálně do všech stran)
- východ (posunutí levé hranice horizontálně).



Obr. 9 Vizualizace BorderLayout rozvržení [vlastní]

Do středu aplikace jsem umístil panel se dvěma záložkami – pro stylovaný text a pro text se značkami. Při zapnutí aplikace v normálním chodu, bude zobrazen panel, ve kterém je možné stylovat text. Panel se značkami obalující stylovaný text není v aplikaci tím hlavním pro uživatele, tedy bude přístupný, ale ne viditelný po celou dobu.

V severní části rozvržení se nachází hlavní ovládací prvky – menu a panel s nástroji. V menu jsou umístěné 3 nabídky: Soubor, Úpravy a Nastavení.

Při vytváření prvků na okně jsem vytvořil několik Action tříd. Každá Action třída bude zodpovědná za určité nastavení. Třídy akcí jsem vytvářel hlavně z důvodu, abych mohl jednu akci přiřadit tlačítku na liště nástrojů a druhou položce v menu bez přidávání dvou posluchačů. Díky tomu bych měl získat o něco více přehledný kód.

Příklad Action tříd:

- FileAction třída – zodpovědná za práci se souborem. Načítání, ukládání nebo zavírání souboru. Načítání a ukládání je dovoleno jen do souborů TXT obsahující čistý text. Vstup není dále nijak kontrolován, požadované úpravy textu jsou na samotném uživateli.
- StyleAction třída – zodpovědná za aplikaci předvolených stylů jako centrování textu, ztučnění textu nebo aplikace kurzívy. Aplikaci stylů neprovádí samotná třída, pouze zprostředkuje vyvolání akce.

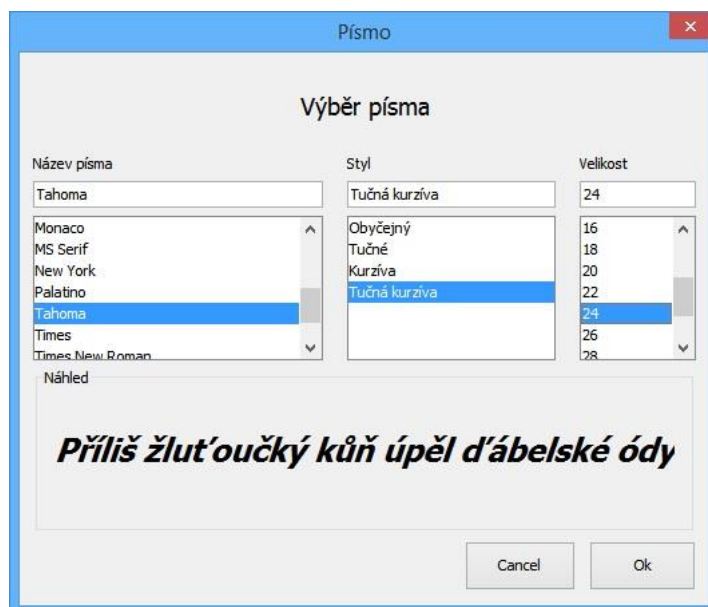
4.2 JDialog pomocná okna

V hlavním okně je možno vyvolat celkem 3 další modální okna, které slouží k dalšímu nastavení komponent a programu. Všechny okna jsou modální, tedy při vyvolání čekají na zadání požadovaných informací, a dokud nedojde k jejich zavření, tak znemožňují práci na okně, které je vyvolalo.

4.2.1 Písmo

Dialog s nastavením písma slouží pro nastavení stylů. Uživatel může volit písmo, jeho styl nebo velikost. Vybrané písmo je vizualizováno v dolní části okna. Do dolní části okna je přidán panel obsahující pangram. Pangram je takový úsek textu, který obsahuje všechna písmena abecedy. Pangram je vhodný, pokud chci vidět zobrazení všech znaků včetně diakritiky.

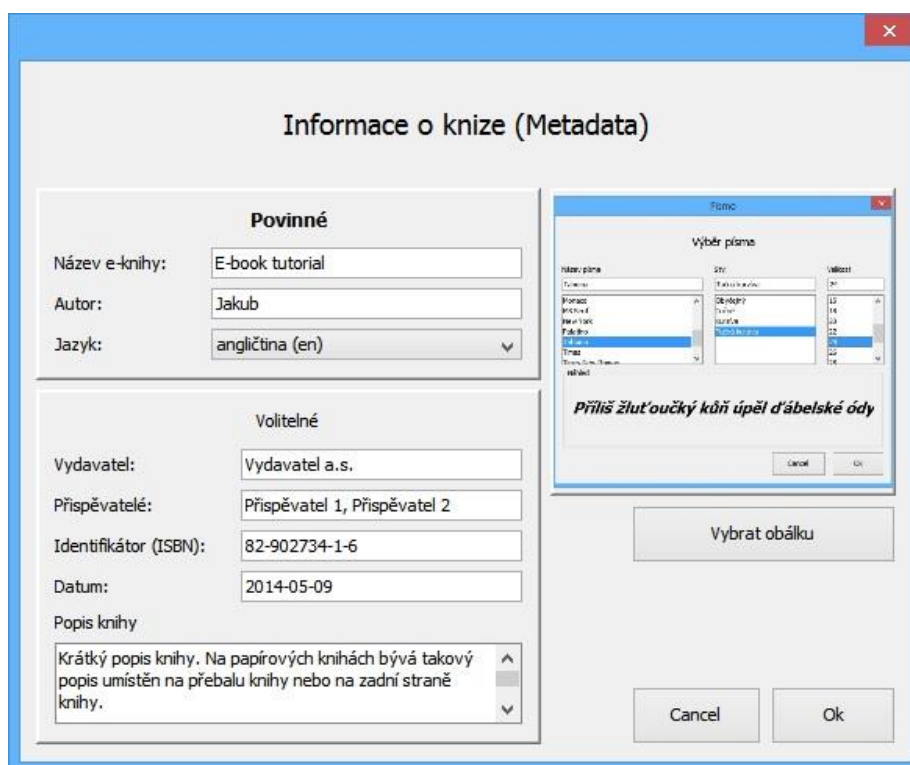
Do pole se styly jsem přidával webově bezpečné styly, u kterých by neměl nastat problém s kódováním. Nedoporučuje se měnit velikost písma zvlášť u každého odstavce. Vzniklý dokument by mohl působit neupraveně a zmateně. Změna písma tedy aplikovaná na celý dokument. [22]



Obr. 10 Okno s nastavením písma [vlastní]

4.2.2 Informace o knize

Okno informace o knize slouží pro nastavení metadat v elektronické knize při překladu. Informace jako název knihy, autor nebo jazyk jsou základní požadované informace v elektronické knize.[22, 34]



Obr. 11 Okno s nastavením informací o knize [vlastní]

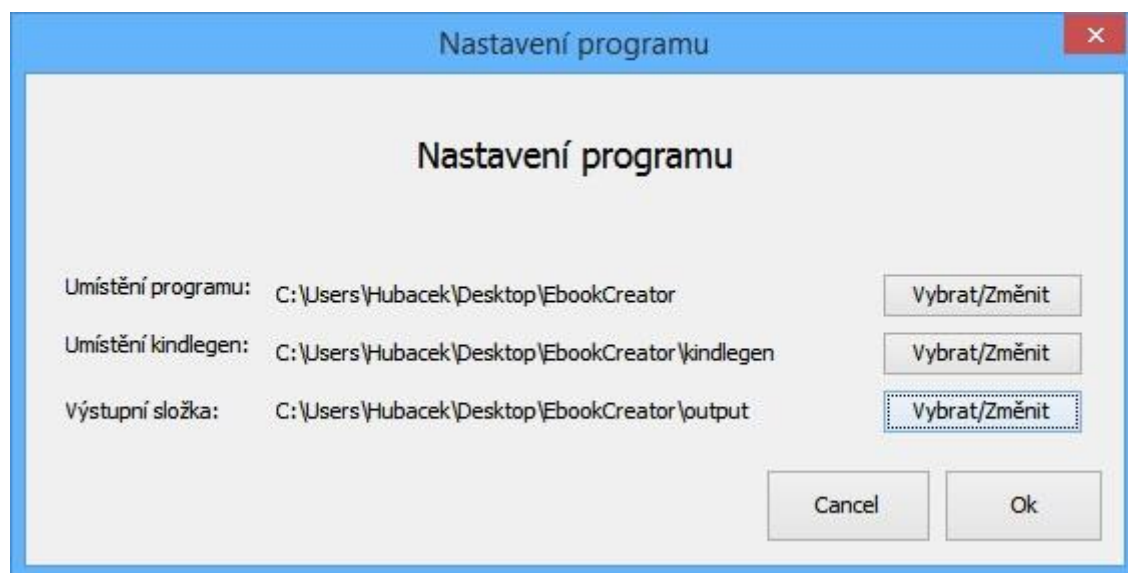
4.2.3 Metadata

Pomocná třída pro skladování informací o knize. Vytvoří se při zavření okna s informacemi o knize. Poté se předává třídě Convertor. Informace o knize jsem rozdělil na povinné údaje, které by měla každá kniha mít, a volitelné, které se u různých publikací nemusí objevovat.

4.2.4 Nastavení programu

Okno s nastavením obsahuje nastavení cest k jednotlivým částím programu. Umístění programu je základní složka, ve které bude výsledný program pracovat. Pod umístěným nástrojem KindleGen by měl mít uživatel přímo složku s nástrojem KindleGen, bez něj totiž nemůže dojít k překladu. Výstupní složka bude lokace, ve které se po překladu objeví hotová e-kniha.

Při prvním spuštění okna Nastavení programu se nastaví aktuální adresář, ve kterém program je. Pokud je složka s programem umístěna na ploše bude vše vypadat jako v ukázce níže.



Obr. 12 Okno s nastavením cest [vlastní]

4.3 StyleManager

StyleManager je třída, která zodpovídá za nastavení stylů v textové oblasti hlavního okna aplikace.

Metody pro získávání vlastností z StyledDocument:

- Style addStyle(String nm, Style parent);
 - Přidá do dokumentu nový styl (například kurzívu).
- Style getStyle(String nm);

- Získám z dokumentu hledaný styl
- `setCharacterAttributes(int offset, int length, AttributeSet s, boolean replace);`
 - Nastaví `AttributeSet` na nastavený text. Je možné nahradit všechny atributy ovlivňovaného textu nebo nové a staré sloučit.
- `Element getCharacterElement(int pos);`
 - Získám element na dané pozici. Z elementu můžu vyčíst aplikované styly.

4.4 TagManager

Třída `TagManager` obsahuje v podstatě jen lineární seznam, který ukládá pozice jednotlivých značek. Jelikož mám část aplikace, kde měním vizualizaci textu a druhou část aplikace, kde měním samotný obsah textu, musím vytvořit správce přidávaných značek.

Lineární seznam je tvořený ze značek. Každá značka obsahuje textovou část a offset od předchozí značky. V případě první značky bude offset vzdáleností od začátku dokumentu.

Při každé aplikaci dalšího stylu, díky správci značek vím, o jaký offset posunu další značku v textu, aby odpovídala skutečně vybranému textu.

4.5 Zpracování XML souboru

XML je rozšířitelný značkovací jazyk (`Extensible Markup Language`), vyvinut a standardizován konsorciem `World-Wide Web Consortium (W3C)`. Jazyk XML nepatří mezi programovací jazyky, nepopisuje žádný sled akcí nebo příkazů. XML vznikl pro serializaci dat podobně jako `JSON`, tedy značí strukturu daného dokumentu. Nedbá na to, jak dokument skutečně vypadá, ale dbá o správu jeho obsahu a to velmi striktně. [30]

Koncovka dokumentu bývá nejčastěji XML, pokud se jedná o standardní zpracování a použití. Takto zpracovaný soubor se používá zejména na uložení dat nebo přenos dat mezi různými počítači.

```
1 <?xml version="1.0" encoding="UTF-8" ?> //XML deklarace
2 <zamestnanci> //začátek kořenového elementu
3   <zamestnanec id="1"> //začátek uzlu s atributem
4     <jmeno>Karel</jmeno> //vnoreny element
5     <prijmeni>Kohák</prijmeni>
6     <fotka>image1.jpg</fotka>
7     <telefon type="mobile">123456789</telefon>
8   </zamestnanec>
9   <zamestnanec id="2"> //začátek dalšího uzlu zamestnanec
10     <jmeno>Sam</jmeno>
11     <prijmeni>Sudý</prijmeni>
12     <fotka>image2.jpg</fotka>
13     <telefon type="home">666789456</telefon>
14   </zamestnanec>
15 </zamestnanci> //konec kořenu zamestnanci
```

Obr. 13 Ukázka dokumentu XML [vlastní]

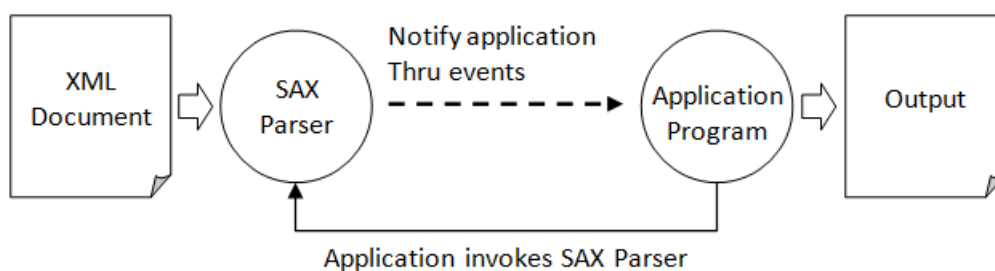
V příkladu vidíme velmi jednoduchou strukturu XML dokumentu. V hlavičce vidíme verzi na XML a kódování. Kódování může být různé, například ISO-8859-1 nebo UTF-8. V našem příkladu je právě UTF-8 tedy bezpečný formát pro českou jazykovou sadu a neměl by s ním být problém.

Celý dokument musí být umístěný v určitém kořenovém elementu. V našem případě se jedná o data o zaměstnancích, proto jsem zvolil element `zamestnanci`. Element `zamestnanci` bude skladovat všechny další zaměstnance. Každý zaměstnanec bude mít svůj vlastní uzel.

Uzel `zamestnanec` v našem příkladu obsahuje data o jednom zaměstnanci: jméno, příjmení, fotka a telefon. Abychom od sebe jednotlivé uzly zaměstnanců odlišili, přidal jsem do elementu `zamestnanec` atribut `id`. Hodnota atributu `id` bude fungovat jako počítadlo, tedy při hledání devátého zaměstnance, hledám zaměstnance s hodnotou atributu `id` 9.

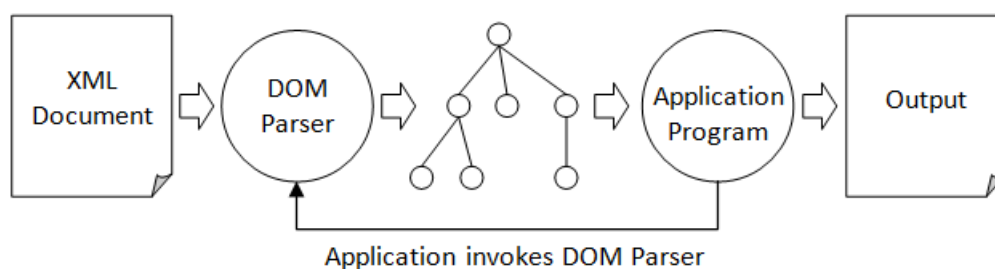
V Javě jsou 2 základní druhy přístupu k XML dokumentům:

- SAX parser: Vyvolává události při průchodu XML dokumentem. Pokud například narazí na tag `<jmeno>` vyvolá například událost `tagJmenoStarted`. Obdobně na konci tagu vyvolá událost oznamující konec tagu. Nepamatuje si celý dokument, ale pouze data, která zpracovávám při obsluze událostí. Přidávání a odebírání elementů je problémové. Kvůli tomu, že nepotřebuje procházet celý strom, je o něco rychlejší a potřebuje daleko méně paměti.



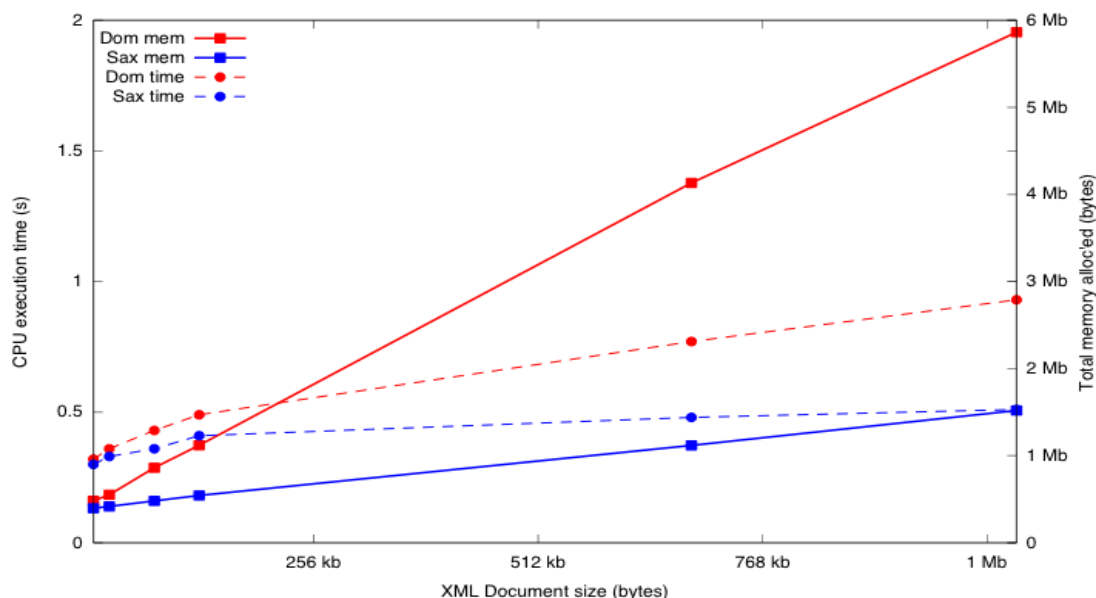
Obr. 14 Vizualizace použití SAX parseru [30]

- DOM parser: Uloží celý XML dokument do paměti při zpracování. Dokument má stromovou strukturu a nemusím se starat o vyvolávání událostí, můžu přidávat atributy a elementy velmi jednoduše. Dokonce můžu jednotlivé větve dávat na jiné místo ve stromové struktuře dokumentu. Pro práci s velkými dokumenty se příliš nehodí, protože pak zabírá i nebezpečně velké množství paměti. Je také o něco pomalejší než SAX parser.



Obr. 15 Vizualizace použití DOM parseru [30]

Pro práci s většími soubory by se tedy měl používat SAX parser, protože DOM může alokovat nebezpečně mnoho paměti. [30]



Graf 4 Porovnání DOM parser a SAX parser [31]

V grafu porovnání vidíme, že při procházení dokumentu vezme DOM asi dvakrát tolik výpočetního času a někdy i pětikrát tolik paměti v porovnání se SAX.

Přesto budu pro práci s daty používat Document Object Model přístup. Přestože je u velkých souborů náročný na paměť a výpočetní čas. Dovoluji si ho použít, protože nebudu tvořit žádné velké XML soubory. Předpokládám, že velikost používaných XML souborů nepřesáhne 30 kB.

```

1. DocumentBuilderFactory docFactory = DocumentBuilderFactory.newInstance();
2. DocumentBuilder docBuilder = docFactory.newDocumentBuilder();
3. Document doc = docBuilder.newDocument();
4.
5. Element rootElement = doc.createElement("knihovna"); //kořenový element
6. doc.appendChild(rootElement);
7.
8. Element kniha = doc.createElement("kniha"); //element pod kořenem
9. rootElement.appendChild(kniha);
10. kniha.appendChild(doc.createTextNode("Kniha o knihách"));
11. kniha.setAttribute("id", "1");
12.
13. TransformerFactory transformerFactory = TransformerFactory.newInstance();
14. Transformer transformer = transformerFactory.newTransformer();
15. DOMSource source = new DOMSource(doc);
16. StreamResult result = new StreamResult(System.out); //output na konzoli
17.
18. transformer.transform(source, result);

```

Obr. 16 Ukázka zdrojového kódu pro tvorbu XML souboru:[32,34, vlastní]

V ukázce zdrojového kódu vytvořím instanci třídy *Document doc*. Tuto instanci budu plnit XML elementy, které budou obsažené v požadovaném výstupu. Přidal jsem element *knihovna*. Tento element je pro mě kořenový, všechny ostatní elementy budou spadat právě pod tento element.

Následně přidávám do elementu *knihovna* element *kniha*. Kniha má parametr *id* s hodnotou 1 a text Kniha o knihách. Element *kniha* může mít další elementy a atributy.

XML výstup

```
1 <?xml version="1.0" encoding="UTF-8" standalone="no"?>
2 <knihovna>
3   <kniha id="1">Kniha o knihách</kniha>
4 </knihovna>
```

Obr. 17 Výstup ukázkového XML generátoru [vlastní]

Ve formátovém XML výstupu vidíme, že kořenový element *knihovna* má vnořený element *kniha*. Element *kniha* má atribut *id* s hodnotou 1 a textovou část Kniha o knihách. Tedy vše proběhlo podle očekávání a vrátilo správný výstup.

4.6 Convertor

Třída Convertor řeší finální část programu. Existuje obsah, který chci uložit do elektronické knihy, a existují vlastnosti, které výsledné e-knize chci předat. Sjednocení všech souborů s obsahem a souborů s vlastnostmi dokumentu je právě úkol třídy Convertor.

Při vytvoření předávám obsah textu se značkami a informace o dokumentu (*metadata*). V daném obsahu postupně nahrazuji značky, značící aplikované styly, konkrétními HTML elementy. Například v případě nadpisu kapitoly, kromě pouhého nahrazení, zjišťuji počet kapitol. To využiji v generování tabulky obsahu. Konce stránek jsou umístěné před každým novým názvem kapitoly v podobě CSS vlastnosti: *page-break-after*. [22, 29]

Po získání upraveného obsahu a vygenerování tabulky obsahu potřebuji ke kompletaci e-knihy ještě XML soubory s vlastnostmi dokumentu. Potřebuji navigační XML soubor (NCX) a soubor s vlastnostmi dokumentu (OPF). O vygenerování se stará třída XML Builder.

4.6.1 XML Builder

Třída XML Builder má za úkol tvorbu všech XML souborů, které při tvorbě e-knihy vzniknou. Jak již bylo zmíněno u popisu XML, budu používat DOM parser. Při vytvoření instance předávám metadata a seznam souborů e-knihy. V seznamu jsou dokumenty seznamy odkazů nebo obrázky.

XML Builder vytvoří 2 části knihy. První je navigační seznam XML dokumentu (NCX soubor) a druhá část je soubor OPF, který obsahuje další informace o vlastnostech a souborech knihy.

4.6.1.1 Popis OPF souboru

```
1 <?xml version="1.0" encoding="utf-8" ?>
2 <package version="2.0" xmlns="http://www.idpf.org/2007/opf" unique-identifier="BookId">
3
4     <!-- Metadata section -->
5     <!-- Manifest section -->
6     <!-- Spine section -->
7     <!-- Guide section -->
8
9 </package>
```

Obr. 18 Ukázka struktury OPF dokumentu [vlastní]

V ukázce OPF dokumentu můžeme vidět, že je nastavené kódování UTF-8 a jsou přidány obecné informace o elementu *package*. Je zde definován *namespace* pod atributem *xmlns*. S definovaným *namespace* můžeme používat další rozšíření dokumentu XML.

```
1 <metadata
2   xmlns:dc="http://purl.org/dc/elements/1.1/"
3   xmlns:opf="http://www.idpf.org/2007/opf">
4
5   <dc:title>Kniha knih</dc:title>
6   <dc:creator opf:role="aut">Jakub Hubáček</dc:creator>
7   <dc:language>cs</dc:language>
8   <dc:identifier id="BookId">57b6-41ee</dc:identifier>
9
10 </metadata>
```

Obr. 19 Ukázka sekce metadata v souboru OPF [vlastní]

Sekce metadata uschovává informace o e-knize. Její autor, její název a další. Je zde přidán další namespace a využívám elementu, například elementu *<dc:title>* pro definování vlastností dokumentu.[22, 29]

```
1 <manifest>
2   <item href="cover.jpg" id="cover" media-type="image/jpeg" />
3   <item href="toc.ncx" id="ncx" media-type="application/x-dtbncx+xml" />
4   <item href="uvod.html" id="uvod" media-type="application/xhtml+xml" />
5   <item href="obsah.html" id="content" media-type="application/xhtml+xml" />
6   <item href="toc.html" id="toc" media-type="application/xhtml+xml" />
7 </manifest>
```

Obr. 20 Ukázka sekce manifest v souboru OPF [vlastní]

V sekci manifest je seznam souborů e-knihy. E-knihy se běžně rozdělují na několik částí: obálka, úvod, text a tabulka obsahu. V sekci manifest definuji naprosto všechny dokumenty, ze kterých je kniha složena bez ohledu na to jestli je uživatel uvidí. [22, 29]

```
1 <spine toc="ncx">
2   <itemref idref="uvod" />
3   <itemref idref="toc" />
4   <itemref idref="obsah" />
5 </spine>
```

Obr. 21 Ukázka sekce spine v souboru OPF [vlastní]

Spine je páteř dokumentu. Definuji pouze dokumenty, které budou uživateli zobrazeny. Hlavním úkolem této páteře e-knihy je definovat posloupnost dokumentů. V ukázce vidíme dokument úvodu následovaný tabulkou odkazů (*Table Of Contents*) a jako poslední jde samotný text knihy. [22, 33]

```
1 <guide>
2   <reference href="toc.html" type="toc" title="Table of Contents" />
3   <reference href="uvod.html" type="text" title="Beginning" />
4 </guide>
```

Obr. 22 Ukázka sekce guide v souboru OPF [vlastní]

Poslední sekci je *guide*, tedy průvodce dokumentu. Například Kindle má možnost se v e-knize kdykoliv přepnout na jinou část knihy. Nejčastěji se pro rychlý přístup definuje právě úvod dokumentu a tabulka s odkazy. [22, 33]

4.6.1.2 Popis NCX souboru

NCX soubor obsahuje XML formát dat představující navigaci v dokumentu elektronické knihy. Zde jsou definovány odkazy na vybrané části dokumentu. Ve výsledku to můžeme považovat za XML rejstřík obsahu e-knihy. Přidáním NCX poskytne o něco více informací o knize i při čtení. [33]


```

1 <?xml version="1.0" encoding="UTF-8" standalone="no"?>
2 <!DOCTYPE ncx PUBLIC "-//NISO//DTD ncx 2005-1//EN"
3 "http://www.daisy.org/z3986/2005/ncx-2005-1.dtd">
4 <ncx
5   xmlns="http://www.daisy.org/z3986/2005/ncx/" version="2005-1" xml:lang="en">
6   <docTitle>
7     <text>kniha</text>
8   </docTitle>
9   <docAuthor>
10    <text>autor</text>
11  </docAuthor>
12  <!--navMap-->
13 </ncx>
14

```

Obr. 23 Ukázka struktury OPF dokumentu [vlastní]

Struktura se příliš neliší od dokumentu OPF. Také zde přidávám namespace a definuji kódování dokumentu. Přidání názvu dokumentu a jméno autora e-knihy zde není přímo povinné. Pokud ale chceme, aby kniha byla jednotná, je dobré mít tyto informace i zde.

```

1 <navMap>
2   <navPoint id="navpoint-1" playOrder="1">
3     <navLabel>
4       <text>Seznam obsahu</text>
5     </navLabel>
6     <content src="../html/toc.html"/>
7   </navPoint>
8   <navPoint id="navpoint-2" playOrder="2">
9     <navLabel>
10      <text>Kapitola 1</text>
11    </navLabel>
12    <content src="../html/content.html#chapter1"/>
13  </navPoint>
14 </navMap>

```

Obr. 24 Ukázka struktury OPF dokumentu [vlastní]

4.6.2 Převod do e-knihy

K převodu potřebuji HTML dokumenty s obsahem, XML dokumenty s informacemi popisující knihu a popřípadě obrázek obálky knihy. Všechny vytvořené soubory si připravuji v pracovním adresáři programu. Obálka knihy, která není vytvořena programem, ale zadaná uživatelem je umístěna do pracovní složky jako kopie původního souboru. [32]

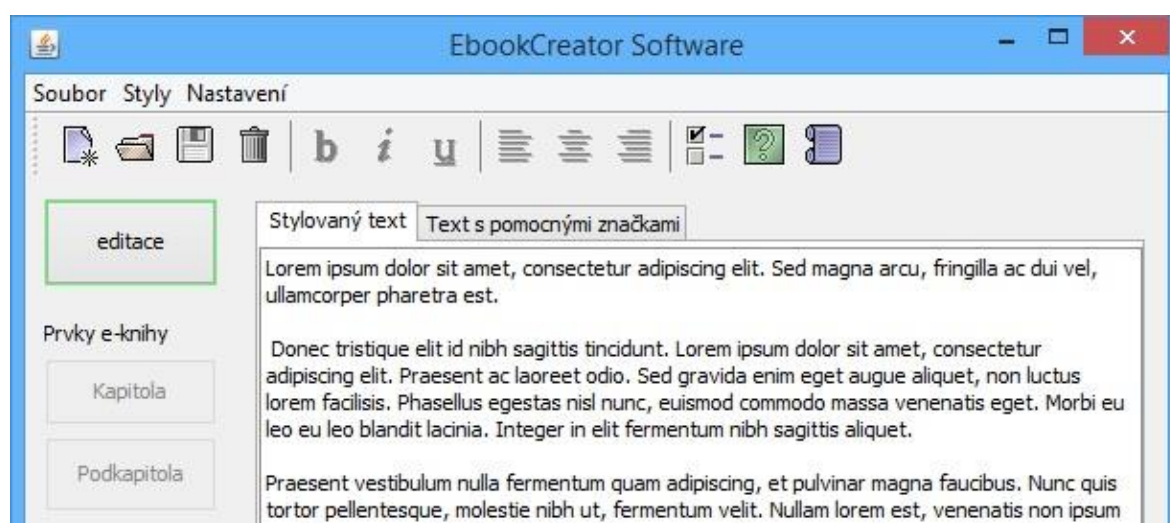
Po umístění všech požadovaných souborů se zavolá nástroj KindleGen. KindleGen na základě odkazů v souboru OPF poskládá celou knihu a navíc zavede kompresi textu. [22]

Vzniklou elektronickou knihu umístím do složky output ve stejném adresáři jako KindleGen a program. Soubory z pracovní složky poté vymažu, čímž připravím program na další překlad.[34]

5 UŽIVATELSKÁ PŘÍRUČKA S PRAKTICKÝMI UKÁZKAMI

Před spuštěním aplikace – zkontrolovat zda se složka s nástrojem KindleGen nachází ve stejném adresáři jako program. Doporučuji rozvržení složky nijak neměnit. Složka s programem by neměla v názvu obsahovat mezery, raději volte spojovací znaky jako pomlčku nebo podtržítko.

5.1 První kroky



Obr. 25 Okno aplikace při spuštění [vlastní]

Při spuštění je aplikace ve stavu, kdy uživatel nemá dostupné editační prvky. Uživatel si musí nejdříve připravit text k editaci a teprve poté může začít aplikovat styly. Stisknutím tlačítka editace se připraví text k úpravám a odemknou se možnosti.

Orientace v okně:

- Menu – obsahuje nabídky pro otevření souboru, uložení souboru (ukládá do formátu txt) nebo vyčištění pracovních dat. Další menu jako styly a možnosti obsahují stejné akce jako panel nástrojů, navíc zde vidět ke každé akci zkratku a popis.
- Panel nástrojů – jsou zde umístěné základní prvky pro stylování dokumentu, jak pro editaci znaků (tučné, kurzíva, podržené) tak pro stylování odstavců (zarovnání vpravo, vlevo a na střed)
- Panel vlevo – obsahuje položky charakteristické pro e-knihy, jako je kapitola a podkapitola, ty jsou určeny pro stylování nadpisů kapitol nebo podkapitol.

- Panel s textem – zde je pracovní plocha, ve které se bude stylovat a upravovat dokument. Druhý panel aplikace s názvem Text s pomocnými značkami obsahuje text obalený značkami na základě formátování zvoleného v panelu Stylovaný text.



Obr. 26 Ukázka aplikovaných stylů [vlastní]

V ukázce jsem aplikoval na slova „Kapitola 1“ a „Kapitola 2“ styl kapitoly. Text v prvním odstavci je centrován doprava a text v druhém odstavci je centrován na střed. Pro program a převod doporučuji, aby se styly na jenom místě aplikovaly pouze jednou. Při aplikování více stylů na jeden prvek můžu výstupní formát vypadat nejednotně.

Aplikací kapitoly uživatel zároveň nastavuje, že se přesně před označeným názvem kapitoly zalomí stránka e-knihy. Na základě aplikovaných kapitol se při převodu vytváří tabulka obsahu.

5.2 Menu

5.2.1 Informace o knize

V menu informace o knize nastavuje uživatel metadata knihy. Může vybrat obrázek, který se umístí jako obálka knihy. Povinné informace jsou jméno autora, jazyk a název knihy.

Pokud požadovaný jazyk není dostupný, vyberte položku „jiný“.

Pokud do e-knihy přispívalo více lidí, oddělte je jejich jména čárkou.

5.2.2 Nastavení programu

Menu nastavení programu uchovává cesty k důležitým adresářům. Jelikož finální převod zprostředkuje KindleGen je potřeba lokalizovat jeho složku. Standardně je KindleGen součástí balíku s programem a tedy cesta by měla být nastavena správně. Tyto údaje většinou není potřeba měnit.

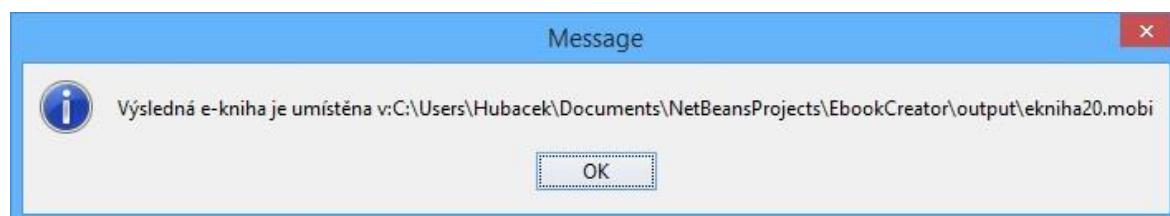
5.2.3 Nastavení písma

Menu nastavení písma je nabídka, která bude do budoucna ještě využita. Standart elektronických nedoporučuje mít více velikostí písma v aplikaci. O zvětšování a zmenšování fontů se většinou starají sami čtečky e-knih.

5.3 Převod knihy

Pokud máte vše nastavené, jak potřebujete, následuje finální fáze. V nabídce soubor je možnost Uložit e-book. Po vybrání této možnosti se začne kompletovat celá kniha, všechny soubory jako OPT, NCX nebo TOC se bude snažit program vytvořit.

Při úspěšném provedení zobrazí například:



Obr. 27 Informace o úspěšném převodu knihy [vlastní]

Do výstupní složky, definované v menu Nastavení programu, se uloží výsledný soubor. Na jménu příliš nezáleží. Čtečky e-knih se nezajímají o název samotného souboru, ale o jeho metadata.

6 PUBLIKACE PRÁCE A JEJÍ BUDOUCNOST

První verze výsledného programu ve výsledku nabízí docela schopné převedení prostého textu do formátu elektronické knihy. Nabídnutím programu komunitě zjistím zpětnou vazbu hodící se pro další vývoj programu.

Výsledný program umístím na 3 fóra zabývající se elektronickými knihami. Pokud má výsledek práce něco přinést, tak by měl být nabídnut komunitě lidí, co mají k e-knihám blízko.

Doufám, že komunita přijde s kritikou programu a tím nastaví nové možnosti vývoje programu. Pokud se umístění na fórech neosvědčí, jako dobrý způsob pro získání zpětné vazby, můžu zkusit přejít na jiné fóra nebo zkusit program nabídnout na zahraničních diskusních fórech.

Umístěno na internetových fórech:

- EbookForum, vše o elektronických knihách
 - Dostupné: <http://www.ebookforum.sk/viewtopic.php?f=90&t=7526>
- Ebook-ctecky.cz, Fórum pro fanoušky ebook čteček
 - Dostupné: <http://www.ebook-ctecky.cz/forum/viewtopic.php?f=13&t=668>
- E-book fórum - knihy na internetu
 - Dostupné: <http://forum.ebookeater.cz/viewtopic.php?f=47&t=23269>

Program pro internetové fóra je dostupný na adrese:

- <http://jakubhubacek.cz/EbookCreator.rar>

6.1 Další vývoj

V další verzi nebo verzích aplikace se chci zaměřit na rozšíření hlavního okna, současná nabídka poskytuje základní prvky. Kromě základních prvků se chci zaměřit i na další typy souborů, které lze přidat do elektronické knihy jako – obrázky, tabulky nebo vektorové formáty obsahující grafiku. Některé typy můžou být zbytečné a nebude důvod je používat, jiné můžou, nabídnou zajímavé oživení.

Současná verze aplikace nepodporuje cizí jazyky. V IDE NetBeans je velice zajímavý nástroj na přidání vícejazyčné podpory aplikace. Chci zjistit, jak moc je vhodný pro internacionalizaci a popřípadě ho využít. Nejdůležitější cizí jazyk pro mě bude angličtina. S projektem přeloženým do angličtiny bych mohl získat další cenou kritiku na zahraničních fórech zaměřených na tematiku elektronických knih.

Přestože jsem navrhl dialogové okno Písmo, nepodařilo se mi najít způsob jak ho správně využít v aplikaci. V budoucnu bych mohl přijít s takovým nastavením písma, kde bude výsledek aplikovaného písma odpovídat, jak standardům, tak přání uživatele.

V souborech MOBI je také možné použít audio a video záznamy; tím může pouhý formát e-knih získat docela jiné využití jako příručky nebo encyklopedie. Největší přípustná velikost všech zkombinovaných souborů je 600 megabajtů.

Při výstupu aplikace se příliš nezajímám o výstup nástroje KindleGen. V dalších verzích aplikace bych se chtěl věnovat i interpretaci výsledků převodu nástroje KindleGen.

ZÁVĚR

V této bakalářské práci jsem navrhl a realizoval aplikaci pro tvorbu elektronických knih z čistého neformátovaného textu do formátu elektronické knihy. Výstupní formát dodržuje aktuální standardy elektronických knih.

V teoretické části jsem popsal historii elektronických knih a čteček od prapočátků do dnešních dní. Rozvedl jsem principy aplikované u čteček elektronických knih, jako snímání textu pomocí ORC nebo displeje čteček na bázi elektronického inkoustu. Uvedl jsem rozdíly jednotlivých formátů elektronických knih a popsal jejich primární použití. V závěru teoretické části se nachází srovnání dostupných řešení převodu textu do elektronické knihy a uvedl mé řešení a jeho výhody.

V praktické části jsem se věnoval návrhu aplikace a následně její realizaci. V návrhu jsem dbal na oddělení částí aplikace a správného předávání informací o knize. Realizace probíhala v IDE NetBeans a aplikace byla naprogramována v jazyce Java. Uvedl jsem ukázky práce s XML soubory v Javě a využití těchto XML souborů pro tvorbu e-knih. IDE NetBeans mi velmi vyhovovalo při tvorbě grafické uživatelského rozhraní i při kompletaci kódu.

K realizovanému programu jsem vytvořil uživatelskou příručku s informacemi o jeho spuštění a tvorbě knihy. Hlavní výhodou mé aplikace je, že uživatel při práci s programem nemusí mít znalosti HTML ani XML a dokáže vytvořit standardizovanou e-knihu. Díky tomu i čtenář bez technických dovedností si dokáže sám vytvořit e-knihu s obálkou, metadaty a seznamem odkazů.

Hotový program s uživatelskou příručkou jsem poskytl komunitám čtenářů e-knih pomocí diskusních fór. Kritika a připomínky z diskusních skupin budou sloužit pro další vývoj programu.

Díky této bakalářské práci jsem si rozšířil znalosti o programovacím jazyku Java a jeho možnostech. Vývoj aplikace mě velice zaujal a budu dál tomuto tématu věnovat. Výsledný program přináší řešení zejména pro uživatele neznalé technických stránek e-knih, pro ty nabízí ideální způsob jak vytvořit e-knihu bez kombinace několika různých nástrojů.

SEZNAM POUŽITÉ LITERATURY

- [1] How Much Does An E-Book Weigh?. In: SMITH, Dave. *International Business Times* [online]. 2011 [cit. 2014-06-03]. Dostupné z: <http://www.ibtimes.com/reporters/dave-smith>
- [2] Getting started with ebook readers. *Ebook Advocate* [online]. 2009 [cit. 2014-06-03]. Dostupné z: <http://www.ebookadvocate.com/?p=8>
- [3] Lessico Tomistico Biculturale. *The Index Thomisticus Treebank* [online]. b.r [cit. 2014-06-03]. Dostupné z: <http://itreebank.marginalia.it/>
- [4] Project Gutenberg. *The History and Philosophy of Project Gutenberg by Michael Hart* [online]. 1992 [cit. 2014-06-03]. Dostupné z: http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart
- [5] DEKHTYAR, Alexander. Ebook Nostalgia. *Work for Idle Hands* [online]. 2010 [cit. 2014-06-03]. Dostupné z: <http://work4idlehands.blogspot.cz/2010/04/ebook-nostalgia.html>
- [6] EBooks: 1998 – The first ebook readers. *Project Gutenberg News* [online]. 2011 [cit. 2014-06-03]. Dostupné z: <http://www.gutenbergnews.org/20110716/ebooks-1998-the-first-ebook-readers/>
- [7] ROBINSON, Greg. Amazon Kindle Review. In: *Big Picture Big Sound* [online]. 2009 [cit. 2014-06-03]. Dostupné z: <http://www.bigpicturebigsound.com/Amazon-Kindle-Review.shtml>
- [8] KOZLOWSKI, Michael. A brief history of eBooks. In: *GOOD E READER* [online]. 2010 [cit. 2014-06-03]. Dostupné z: <http://goodereader.com/blog/electronic-readers/a-brief-history-of-ebooks>
- [9] POLESNÝ, David. Čtečka od eReading.cz umí česky, má české knihy a jde proti Kindlu. *Zive.cz* [online]. 2011 [cit. 2014-06-03]. Dostupné z: <http://www.zive.cz/bleskovky/ctecka-od-ereadingcz-umi-cesky-ma-ceske-knihy-a-jde-proti-kindlu/sc-4-a-159555/default.aspx>
- [10] *EReading: Projekt portálu elektronických knih* [online]. 2010 [cit. 2014-06-03]. Dostupné z: <http://www.ereading.cz/cs/ctecka-knih-ereading-start-2>
- [11] Co je DRM?. In: *Databook.cz* [online]. [cit. 2014-06-04]. Dostupné z: <http://www.databook.cz/DRM>
- [12] KRYBUS, Vojtěch. *Technologie elektronického inkoustu - historie a vývoj*. Brno, 2003. Dostupné z: <http://www.fi.muni.cz/usr/jkucera/pv109/2003/xkrybus2003.htm>. Kolokviální práce. Masarykova univerzita v Brně, Fakulta informatiky.
- [13] WOODFORD, Chris. Optical character recognition. In: *Explain that Stuff* [online]. 2013 [cit. 2014-06-04]. Dostupné z: <http://www.explainthatstuff.com/how-ocr-works.html>

- [14] Ebook Formats: A Quick Guide For Self-Publishers. In: *Bubblecow.net* [online]. 2012 [cit. 2014-06-04]. Dostupné z: <http://bubblecow.net/a-quick-guide-for-self-publishers-worried-about-ebook-formats/>
- [15] Adobe PDF 101 — Quick overview of PDF file format. *Adobe* [online]. 2009 [cit. 2014-06-10]. Dostupné z: http://partners.adobe.com/public/developer/tips/topic_tip31.html
- [16] KASÍK, Pavel. TEST: České čtečky knih. Nejhorší je Wooky, Cybook stojí za úvahu. *Tech-net.cz* [online]. 2012 [cit. 2014-06-04]. Dostupné z: http://technet.idnes.cz/ceske-e-inkove-ctecy-recenze-cybook-odyssey-od-palmknihy-cz-wooky-1dn-tec_technika.aspx?c=A120730_140748_tec_technika_pka
- [17] CLARK. The 6 Most Popular eBook Formats To Know About. *Edudemic.com* [online]. 2013 [cit. 2014-06-04]. Dostupné z: <http://www.edudemic.com/most-popular-ebook-formats/>
- [18] EBook Collections Usage Guide. In: *J. Willard Marriott Library* [online]. 2013 [cit. 2014-06-04]. Dostupné z: <http://campusguides.lib.utah.edu/content.php?pid=163999&sid=1395379>
- [19] SVAZ ČESKÝCH KNIHKUPCŮ. *Zpráva o českém knižním trhu 2012/2013* [online]. 2013 [cit. 2014-05-29]. ISBN 978-80-902495-8-5. Dostupné z: <http://sckn.cz/content/zpravy/file-936.pdf>
- [20] VLČEK, Jiří. Proč není dobré půjčování e-knih. *Palmknihy.cz* [online]. 2013 [cit. 2014-06-04]. Dostupné z: <http://blog.palmknihy.cz/2013/10/proc-neni-dobre-pujcovani-e-knih.html>
- [21] KRČMÁŘ, Petr. Kolik stojí výroba knihy?. *Root.cz* [online]. 2010 [cit. 2014-06-04]. Dostupné z: <http://petrkrmar.blog.root.cz/2010/03/04/kolik-stoji-vyroba-knihy/>
- [22] *Amazon Kindle Publishing Guidelines* [online]. Amazon.com, Inc., 2013 [cit. 2014-01-30]. Dostupné z: <http://kindlegen.s3.amazonaws.com/AmazonKindlePublishingGuidelines.pdf>
- [23] Word. *Microsoft Office* [online]. 2014 [cit. 2014-06-10]. Dostupné z: <http://office.microsoft.com/en-us/word/FX100487981033.aspx>
- [24] GREENE, Araby. Create an NCX File. *CJs Easy As Pie Kindle Tutorial* [online]. [cit. 2014-06-11]. Dostupné z: <http://www.cjs-easy-as-pie.com/p/create-ncx-file-by-araby-greene.html>
- [25] LibreOffice v češtině. *LibreOffice* [online]. 2013 [cit. 2014-06-11]. Dostupné z: <https://cs.libreoffice.org/>
- [26] NESBITT, Scott. Creating an ebook with LibreOffice Writer. *Ubuntu Musings* [online]. 2011 [cit. 2014-06-11]. Dostupné z: <http://scottnesbitt.net/ubuntublog/creating-a-ebook-with-libreoffice-writer/>
- [27] *Mobipocket.com* [online]. 2014 [cit. 2014-06-11]. Dostupné z: <http://www.mobipocket.com/>
- [28] Sigil. *Google Code* [online]. 2014 [cit. 2014-06-11]. Dostupné z: <https://code.google.com/p/sigil/>

- [29] KINDLE DIRECT PUBLISHING. *Building Your Book for Kindle*. Seattle: Amazon Digital Services, Inc., 2012. ISBN ASIN B007URVZJ6.
- [30] *Java Programming, Java & XML* [online]. 2009 [cit. 2014-06-12]. Dostupné z: http://www.ntu.edu.sg/home/ehchua/programming/java/j6d_xml.html
- [31] GORTER. Java XML Parsing: SAX vs DOM. *TECH.inhelsinki.nl* [online]. 2007 [cit. 2014-06-12]. Dostupné z: <http://tech.inhelsinki.nl/2007-08-29/>
- [32] KISZKA, Bogdan. *1001 tipů a triků pro programování v jazyce Java: nejbohatší sbírka řešených problémů a užitečných nápadů v Javě*. Vyd. 1. Brno: Computer Press, 2003, 519 s. ISBN 80-722-6989-5.
- [33] *BB eBooks* [online]. 2013 [cit. 2014-02-05]. Dostupné z: <http://bbebooksthailand.com/index.html>
- [34] DARWIN, Ian F. *Java: kuchařka programátora*. Vyd. 1. Brno: Computer Press, 2006, 798 s. ISBN 80-251-0944-5.

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

| | |
|------|--------------------------------------|
| AZW | Amazon Word. |
| CD | Compact Disc. |
| CSS | Cascading Style Sheets. |
| ČR | Česká republika. |
| DOC | Microsoft Word Document. |
| DOCX | Microsoft Word Document Extended. |
| DPH | Daň z přidané hodnoty. |
| DRM | Digital Rights Management. |
| DVD | Digital Versatile Disc. |
| EPUB | Electronic publication. |
| HTML | HyperText Markup Language. |
| IBM | International Business Machines. |
| IMG | Image. |
| IP | Internet Protocol. |
| JSON | JavaScript Object Notation. |
| LCD | Liquid Crystal Display. |
| LIT | Literature (Microsoft ebook format). |
| MOBI | Mobipocket ebook, |
| NCX | Navigational Center for XML, |
| OCR | Optical Character Recognition. |
| ODT | Open Document Text. |
| OPF | Open Packaging Format. |
| PC | Personal Computer. |
| PDA | Personal Digital Assistant. |

| | |
|-----|-----------------------------|
| PDB | Palm Database. |
| PDF | Portable Document Format. |
| TXT | Text. |
| W3W | Worl-Wide Web Consortium. |
| XML | Extensible Markup Language. |

SEZNAM OBRÁZKŮ

| | |
|--|----|
| Obr. 1 Zleva Sony PRS-600, Amazon Kindle 2 a Barnes & Noble Nook [2]..... | 10 |
| Obr. 2 Jedna z prvních čteček - čtečka RCA [5] | 12 |
| Obr. 3 Amazon Kindle 2. generace [7] | 12 |
| Obr. 4 Čtečka eReading.cz [10]..... | 13 |
| Obr. 5 Navigační body (Tick Marks) v čtečce Kindle [24] | 21 |
| Obr. 6 Okno s nastavením informací v programu Writer [26] | 22 |
| Obr. 7 Editace knihy v aplikaci Sigil [28] | 23 |
| Obr. 8 Vizualizace práce programu [vlastní]..... | 25 |
| Obr. 9 Vizualizace BorderLayout rozvržení [vlastní] | 28 |
| Obr. 10 Okno s nastavením písma [vlastní]..... | 30 |
| Obr. 11 Okno s nastavením informací o knize [vlastní] | 30 |
| Obr. 12 Okno s nastavením cest [vlastní] | 31 |
| Obr. 13 Ukázka dokumentu XML [vlastní]..... | 33 |
| Obr. 14 Vizualizace použití SAX parseru [30]..... | 34 |
| Obr. 15 Vizualizace použití DOM parseru [30] | 34 |
| Obr. 16 Ukázka zdrojového kódu pro tvorbu XML souboru:[32,34, vlastní] | 35 |
| Obr. 17 Výstup ukázkového XML generátoru [vlastní] | 36 |
| Obr. 18 Ukázka struktury OPF dokumentu [vlastní] | 37 |
| Obr. 19 Ukázka sekce metadata v souboru OPF [vlastní] | 37 |
| Obr. 20 Ukázka sekce manifest v souboru OPF [vlastní]..... | 38 |
| Obr. 21 Ukázka sekce spine v souboru OPF [vlastní] | 38 |
| Obr. 22 Ukázka sekce guide v souboru OPF [vlastní]..... | 38 |
| Obr. 23 Ukázka struktury OPF dokumentu [vlastní]..... | 39 |
| Obr. 24 Ukázka struktury OPF dokumentu [vlastní]..... | 39 |
| Obr. 25 Okno aplikace při spuštění [vlastní] | 40 |
| Obr. 26 Ukázka aplikovaných stylů [vlastní] | 41 |
| Obr. 27 Informace o úspěšném převodu knihy [vlastní] | 42 |

SEZNAM GRAFŮ

| | |
|---|----|
| Graf 1 Odhad počtu prodaných e-knih v ČR [19] | 18 |
| Graf 2 Odhad počtu dostupných titulů ke konci roku [19] | 19 |
| Graf 3 Prodejnost e-knih ve vybraných žánrech [19] | 19 |
| Graf 4 Porovnání DOM parser a SAX parser [31] | 35 |

SEZNAM PŘÍLOH

Software pro tvorbu e-knih