

Univerzita Karlova

Pedagogická fakulta

Katedra informačních technologií a technické výchovy

## **BAKALÁŘSKÁ PRÁCE**

**System pro generování statického webu**

**System for static web generation**

Emil Miler

Vedoucí bakalářské práce: PhDr. Josef Procházka, Ph.D.

Studijní program: Specializace v pedagogice

Studijní obor: Informační technologie se zaměřením na vzdělávání

Praha 2020

Odevzdáním této bakalářské práce na téma Systém pro generování statického webu potvrzuji, že jsem ji vypracoval pod vedením vedoucího práce samostatně za použití v práci uvedených pramenů a literatury. Dále potvrzuji, že tato práce nebyla využita k získání jiného nebo stejného titulu.

Praha, 4. května 2020

Podpis autora

Děkuji především svému vedoucímu PhDr. Josefu Procházkovi, Ph.D. za všestrannou podporu během všech fází zpracování této práce, především pak za jeho vstřícnost, trpělivost a pochopení.

Mou povinností je také poděkovat Lukáši Hozdovi za jeho technickou a odbornou asistenci se statickými generátory a se sázením samotné práce, a dále také Bc. Janu Vaisovi za jeho cenné rady k sázení a k obecnému psaní závěrečných prací.

Dále musím poděkovat Vítu Matějčkoví, Albertu Pospíšilovi a Dawidu Kubišovi za pomoc s překlady, korekturou a za obecné rady k obsahu. Také děkuji Ing. Emilu K. Mironovovi, Ph.D. za pomoc s formální úpravou.

Na závěr děkuji partnerce a rodině za podporu během psaní této práce.

## **Abstrakt**

Bakalářská práce pojednává o problematice statických webových stránek a jejich generátorů. Popisuje nejčastěji používaná webová paradigmatata, generátory a značkovací jazyky, které se ke psaní obsahu těchto stránek používají. Součástí práce je i taxonomie požadavků a jejich modelová implementace ve formě celého systému pro generování statických webových stránek, včetně optimalizace a implementace vlastních rozšíření a verzovacího systému pro správu obsahu. V závěrečné části je modelová implementace vyhodnocena na základě dříve zmíněných požadavků a poznatků z předchozí části práce.

## **Klíčová slova**

www, web, statický, generátor

## **Abstract**

This bachelor's thesis deals with static websites and static web generators. It describes the most used website paradigms, static web generators and typesetting languages, which are used as source files for the content of the website. A part of the thesis is dedicated to the taxonomy of requirements for a specific website and its actual implementation with optimizations, implementations of custom extensions and a Content Management System. The last part of the thesis evaluates the created implementation according to the given requirements and findings from the first part.

## **Keywords**

www, web, static, generator

# Obsah

<b>1 Úvod</b>	<b>7</b>
Úvod	7
1.1 Cíle práce . . . . .	7
1.2 Struktura práce . . . . .	8
<b>2 Staticky generované webové stránky</b>	<b>9</b>
2.1 Výhody statických webových stránek . . . . .	9
2.2 Princip generátorů . . . . .	12
<b>3 Webová paradigmata</b>	<b>13</b>
3.1 Webová prezentace . . . . .	13
3.2 Index všeobecných informací . . . . .	14
3.3 Technická dokumentace . . . . .	15
3.4 Galerie . . . . .	15
<b>4 Značkovací jazyky</b>	<b>17</b>
4.1 Principy značkovacích jazyků . . . . .	17
4.2 Nejběžnější jazyky . . . . .	18
4.2.1 $\text{\TeX}$ . . . . .	18
4.2.2 Troff . . . . .	19
4.2.3 Markdown . . . . .	20
4.2.4 Org-mode . . . . .	21
4.2.5 reStructuredText . . . . .	22

<b>5</b>	<b>Taxonomie požadavků pro modelový web</b>	<b>23</b>
5.1	Obecná kritéria . . . . .	23
5.2	Kritéria specifická pro modelový web . . . . .	24
5.3	Kritéria pro šablony a design . . . . .	24
<b>6</b>	<b>Modelová implementace</b>	<b>25</b>
6.1	Výběr vhodného systému . . . . .	25
6.1.1	Verzovací systém pro správu obsahu . . . . .	25
6.1.2	Generátor statického webu . . . . .	25
6.2	Tvorba šablony . . . . .	26
6.3	Automatické generování vícevrstvé navigace . . . . .	30
6.4	Rozšíření šablony . . . . .	33
6.5	Optimalizace . . . . .	38
6.5.1	Typy a kvalita obrázků . . . . .	38
6.5.2	Ikona <i>favicon.ico</i> . . . . .	39
6.5.3	Obecné HTML optimalizace . . . . .	39
6.5.4	Optimalizace videa . . . . .	40
6.6	Správa obsahu a verzování . . . . .	41
6.6.1	Automatizace generování obsahu . . . . .	42
<b>7</b>	<b>Vyhodnocení modelové implementace</b>	<b>45</b>
7.1	Návrhy pro rozšíření systému . . . . .	45
7.2	Vyhodnocení implementace vlastních rozšíření . . . . .	46
	<b>Závěr</b>	<b>47</b>
	<b>Seznam použité literatury</b>	<b>48</b>

# 1. Úvod

Od statických stránek, které se používaly při vzniku WWW<sup>1</sup>, se trend posunul ke složitým webovým aplikacím a frameworkům, jako jsou například Wordpress, Joomla, nebo Drupal. Dnes se opět trend obrací směrem ke tvorbě statických webových stránek, které jsou mnohem lepší z hlediska rychlosti a bezpečnosti.

Práce se zabývá problematikou generátorů statických webových stránek. Za statické webové stránky se považují stránky, jejichž obsah je předem vygenerován do HTML<sup>2</sup> pomocí daného generátoru, tedy programu, který z obsahu napsaného v libovolném značkovacím jazyce vygeneruje validní webovou stránku.

## 1.1 Cíle práce

Hlavním cílem práce je vysvětlit čtenáři fungování a sestavení automatizovaného systému pro generování statických webových stránek. Čtenář porozumí problematice statických webových stránek, jejich generátorů, značkovacích jazyků a implementaci těchto konceptů prakticky v modelové implementaci. Tento hlavní cíl je dále rozpracován cíli dílčími:

- seznámit čtenáře se systémy pro generování statických webů,
- analyzovat základní funkce mapovaných systémů,
- sestavit taxonomii požadavků na systém pro generování statického webu,
- navrhnout modelovou implementaci s užitím vybraného systému,
- vyhodnotit funkce daného systému a navrhnout jeho rozšíření.

---

<sup>1</sup>World Wide Web – Celosvětová síť pro prohlížení, ukládání a sdílení dokumentů

<sup>2</sup>Hypertext Markup Language – Sázecí jazyk používaný k popisu obsahu a struktury webové stránky

## 1.2 Struktura práce

Struktura práce je koncipována do čtyř základních částí; v první části je pojednáváno o konceptu statických webových stránek, jejich výhodách a o roli statických generátorů. Další část navazuje analýzou nejčastějších webových paradigmat, která se ve světě WWW objevují, a jsou v ní vybrány nejvhodnější systémy pro generování jejich statického obsahu. Dále následuje kapitola pojednávající o nejběžnějších značkovacích jazycích, kterými je popisován obsah pro vybrané generátory.

V další, praktické části jsou definována kritéria požadavků pro modelovou implementaci, která je poté realizována včetně ukázek kódu a popisu celého principu fungování systému. V závěrečné kapitole je modelová implementace vyhodnocena a jsou zde určeny návrhy na její další rozšíření.



## 2. Staticky generované webové stránky

Princip statické webové stránky sahá až ke vzniku WWW, kdy existovaly pouze stránky statické, tedy stránky nevytvářené dynamicky na základě různých proměnných a na uživatelském vstupu. U statických webových stránek dochází k vytvoření HTML kódu ve chvíli, kdy je změněn zdrojový obsah nebo kdy autor ručně spustí generátor. (PC Magazine, 2020)

Dynamické stránky jsou generovány speciálně pro každého uživatele na základě jeho nastavení, různých vstupů, proměnných a dalších vlastností ovlivňujících výslednou podobu či obsah stránky. Ke generování dochází ve chvíli, kdy si uživatel stránku vyžádá, nikoliv předem, jako je tomu u staticky generovaných stránek. (PC Magazine, 2017)

### 2.1 Výhody statických webových stránek

Pro sdílení statického obsahu mezi různé uživatele stačí velmi jednoduchý HTTP<sup>1</sup> server bez jakýchkoliv dalších modulů typu *PHP*, *Python* a dalších systémů<sup>2</sup>, které by obsah dynamicky generovaly například z dat přečtených z databáze nebo z uživatelského vstupu. Na straně serveru tedy nedochází ke zpracování obsahu těsně před jeho odesláním uživateli, čímž se v komunikaci mezi klientem a serverem výrazně snižuje „Time To First Byte“<sup>3</sup>, a tím dochází ke snížení celkové latence. (Hoffman, 2013)

Snížení latence může pozitivně přispět ke spokojenosti uživatelů, což dokazuje nespočet výzkumů na toto téma, například analýza z webového portálu Financial Times, kde se uvádí, že rychlost webové stránky negativně ovlivňuje hloubku jejího užívání, ať už je odezva sebemenší. Jak je zde rovněž uvedeno, data ukazují, že z pohledu uživatelské spokojenosti a finančního dopadu existují jasné a důležité výhody při zrychlení webové

---

<sup>1</sup>Hypertext Transfer Protocol – Internetový protokol sloužící ke komunikaci s WWW servery.

<sup>2</sup>Jedná se o programovací jazyky nejčastěji používané ke tvorbě dynamických webových aplikací.

<sup>3</sup>Time To First Byte – Čas mezi odesláním požadavku a přijutím prvního bajtu dat.

stránky. Z tohoto výzkumu se autoři rozhodli v měsících po vydání analýzy investovat více času do úprav všech aspektů svých nových webových stránek s cílem jejich zrychlení. (Chadburn – Lahav, 2016)

Eliminováním dynamického obsahu se také předchází nevyžádaným vstupům od uživatele, jež mohou být i cílené na prolomení bezpečnostních nedostatků webové aplikace a v některých případech mohou vést k úniku citlivých dat, převzetí kontroly útočníka nad webovou aplikací nebo celým serverem, podstrčení falešných dat uživateli a mnoha dalším běžným útokům. Statický web eliminuje tento problém, jelikož nemá žádný uživatelský vstup.

Sledování a analýze nejčastějších chyb webových aplikací a serverů se věnuje organizace OWASP<sup>4</sup>, která vydává aktualizované seznamy a statistiky. Podle OWASP (2017) se v roce 2017 objevovaly nejčastěji tyto chyby a bezpečnostní nedostatky:

1. injekce,
2. rozbitá autentizace,
3. odhalení citlivých dat,
4. XML External Entities (XXE),
5. nefunkční řízení přístupu,
6. špatná konfigurace zabezpečení,
7. Cross-Site Scripting (XSS),
8. nezabezpečená deserializace,
9. užívání komponent se známými zranitelnostmi,
10. nedostatečné logování a monitorování.

Většina těchto chyb se vztahuje právě k dynamickým webovým aplikacím. Bezpečnost tedy závisí nejen na programátorovi, jenž aplikaci vytváří, ale také na tom, že programovací jazyk je bezpečně implementován. To nelze tvrdit o nejpoužívanějším jazyce

---

<sup>4</sup>The Open Web Application Security Project (<https://owasp.org/>).

PHP, který obsahuje mnoho chyb, viz statistika nalezených bezpečnostních děr na serveru CVE Details (2020a), který se tímto tématem zabývá. Jazyk zároveň nevede programátora k psaní bezpečného kódu, což má za následek nebezpečené aplikace, pokud si autor nedá pozor na správné ošetření vstupů a dalších bezpečnostních aspektů programu.

Podstatným příkladem chybovosti dynamických webů je systém Wordpress, v němž jsou podle serveru CVE Details (2020b) každý rok nalezeny desítky bezpečnostních chyb, přičemž mnoho dalších přibývá s instalací nezabezpečených rozšíření. Například na začátku roku 2020 byla nalezena bezpečnostní chyba v rozšíření využívaném na více než dvě stě tisících webových stránkách, která útočnickům umožňovala smazat obsah databáze (Khandelwal, 2020). Na konci roku 2019 umožnila chyba ve dvou nezabezpečených rozšířeních neautorizované přihlášení k účtu administrátora bez použití hesla (Khandelwal, 2019).

Údržba velkých webových aplikací je často problematická. Kód je nutné udržovat v návaznosti na aktualizace daného jazyka, databázového systému a dalších aspektů. Těmto aktualizacím se z bezpečnostních důvodů nelze vyhýbat. Statický web nemusí udržovat funkční propojení s databázemi a dalšími částmi systému, a je tedy mnohem méně náročný na dlouhodobou údržbu. Při zvolení správného generátoru není nutná ani údržba šablon a celý systém při zachování stejného prostředí nepřestane fungovat. Protože statický generátor nepracuje s uživatelským vstupem, vyhýbá se bezpečnostním chybám a tím i nutným aktualizacím.

Jako každý jiný systém, i statické generátory mají své nevýhody. Hlavním z problémů je, že tvorba a správa obsahu může být pro běžného uživatele náročnější než skrze klasické webové rozhraní s administračním panelem a jednoduchou správou celého obsahu. Pro přidání nebo úpravu obsahu je nutné pracovat s lokálními soubory ve stromové struktuře a při generování je často potřebný zásah do shellu<sup>5</sup>. Tvorba systému pro automatizované generování je také náročnější než instalace některého z běžných CMS<sup>6</sup>. (Cimpanu, 2015)

---

<sup>5</sup>Program pro interpretování příkazů v prostředí příkazové řádky.

<sup>6</sup>Systém pro správu a obsahu (Wordpress, Joomla, Drupal a podobně).

## 2.2 Princip generátorů

Ekosystém generátoru statického obsahu je tvořen ze tří hlavních složek.

První částí jsou soubory šablon, jež popisují rozložení stránky, vizuální vlastnosti, typografii, ale také vstupní a výstupní kódování a formáty. V podstatě definují, jak a kam se bude obsah vkládat.

Druhou částí je obsah samotný, napsaný v některém ze značkovacích jazyků, jež jsou dále popsány v kapitole 4. Obsah bývá strukturován do sekcí a souborů, aby bylo snadné rozlišit, do které části výsledné stránky patří.

Třetí a poslední složkou je samotné jádro generátoru, jež zpracovává obsah, vkládá ho do šablon a generuje statickou webovou stránku.

Většina generátorů zároveň umí pracovat s konfiguračními soubory, pomocí nichž lze nastavit globální chování generátoru. Část z nich také integruje jednoduchý webserver, který umožňuje autorovi náhled výstupních stránek, zatím co tvoří obsah.

(Cimpanu, 2015)

## 3. Webová paradigmata

Ve světě webových stránek se setkáváme se spoustou forem a paradigmat, jež se hodí pro zpracování různých druhů informací. Neexistuje žádné formální zařazení druhů webových stránek do skupin, ovšem některé webové portály se pokouší určit základní druhy webových stránek, které se na Internetu objevují. Na základě těchto portálů a jejich rozřazení do skupin<sup>12</sup>, jež jsou často mířené na specifický obsah, lze vytvořit pět základních paradigmat. Jsou jimi:

- webová prezentace,
- index všeobecných informací,
- technická dokumentace,
- galerie,
- sociální sítě a fóra.

V této práci bylo ke každému z paradigmat vybráno několik existujících generátorů, které jsou vhodné pro generování a správu daného druhu obsahu. Výjimkou je skupina sociálních sítí a fór, kde staticky generovaný obsah není z důvodu často se měnícího obsahu vhodným řešením.

### 3.1 Webová prezentace

Nejbližší původním webovým stránkám z dob vzniku WWW jsou webové prezentace, tedy stránky s jednoduchým obsahem sloužící k předání informací čtenáři, například formou článků. Do této skupiny lze zařadit portfolia, blog, online noviny a časopisy, firemní stránky a podobně, jež se skvěle hodí ke statickému generování obsahu.

Generátory pro tento typ webové stránky jsou často univerzální a je možné aplikovat většinu těchto generátorů i na ostatní formy stránek. Většina existujících generátorů je

---

<sup>1</sup><http://www.xislegraphix.com/website-types.html>

<sup>2</sup><https://www.hostgator.com/blog/popular-types-websites-create/>

zaměřena právě na toto paradigma, a lze z čeho vybírat. Každým dnem vznikají nové implementace a svět staticky generovaných stránek se rapidně mění.

Tabulka 3.1: Vybrané generátory pro webové prezentace

Program	Jazyk	Podporované jazyky obsahu
blogc <sup>3</sup>	C	Vlastní jazyk blogc-source <sup>4</sup> založený na Markdownu
Hugo <sup>5</sup>	Go	Markdown, HTML, AsciiDoc, Pandoc Markdown, reST
Luapress <sup>6</sup>	Lua	Markdown
Nift <sup>7</sup>	C++	Agnostický, tj. Markdown, L <sup>A</sup> T <sub>E</sub> X, HTML, JS, SQL, ...
Pelican <sup>8</sup>	Python	Markdown, reST
ssg <sup>9</sup>	shell	Markdown (lowdown, Markdown.pl)
Zola <sup>10</sup>	Rust	Markdown (CommonMark)

## 3.2 Index všeobecných informací

Za obecného zástupce tohoto druhu stránek lze považovat Wikipedii, která podnítila vznik spousty jiných takzvaných *Wiki systémů* a stránek. Jde o formu stránky, jež obsahuje velké množství informací s různou provázaností. Nejde o lineární dokumentaci projektu, nýbrž o encyklopedii obecných informací.

Jak již bylo řečeno, tuto formu stránek lze implementovat jakýmkoliv generátorem. Existují ale generátory vytvořené přímo pro účel tohoto paradigmatu. Takových generátorů je ovšem málo, jelikož jedním ze základních znaků wiki je to, že její obsah tvoří komunita uživatelů, což nemusí být vždy vhodné pro systém statického generování obsahu. Jedním takovým programem je ikiwiki<sup>11</sup>, který funguje stejně jako ostatní sta-

<sup>3</sup><https://blogc.rgm.io/>

<sup>4</sup><https://blogc.rgm.io/man/blogc-source.7.html>

<sup>5</sup><https://gohugo.io/>

<sup>6</sup><https://luapress.org/>

<sup>7</sup><https://www.staticgen.com/nift/>

<sup>8</sup><https://blog.getpelican.com/>

<sup>9</sup><https://www.romanzolotarev.com/ssg.html>

<sup>10</sup><https://www.getzola.org/>

<sup>11</sup><https://ikiwiki.info/>

tické generátory. Další programy, například MDwiki<sup>12</sup>, zpracovávají Markdown přímo v prohlížeči a HTML obsah tedy není předgenerován na serveru.

Program MDwiki je prezentován jako generátor vytvořený na míru pro wiki stránky, ovšem jeho výstup je velice podobný generátorům z tabulky 3.2.

### 3.3 Technická dokumentace

Od Wiki stránek se technická dokumentace liší organizováním svého obsahu, který je cílený na přesný popis systému či objektu. Obsah je strukturován jako manuál, nebo-li kniha, kde se obsah každé části věnuje danému tématu a různý obsah je strukturován do vlastních knih.

Tabulka 3.2: Vybrané generátory pro technickou dokumentaci

Program	Jazyk	Podporované jazyky obsahu
mdBook <sup>13</sup>	Rust	Markdown
MkDocs <sup>14</sup>	Python	Markdown
Sphinx <sup>15</sup>	Python	reStructuredText

### 3.4 Galerie

V případě galerií se jedná o obsah složený primárně z obrázků nebo videí uspořádaných ve struktuře pro jejich snadnou prezentaci. Mohou sloužit jako foto alba, časové příběhové osy, nebo jako portfolio produktů či fotografií.

Generátory galerií umožňují jednoduše řadit obrázky do skupin a vytvářet k nim popisky. Markup jazyky zde nehrají tak velkou roli jako u ostatních druhů paradigmat, jelikož hlavním obsahem není text, nýbrž samotné obrázky a videa.

---

<sup>12</sup><https://dynamon.github.io/mdwiki/>

<sup>13</sup><https://github.com/rust-lang/mdBook/>

<sup>14</sup><https://www.mkdocs.org/>

<sup>15</sup><https://www.sphinx-doc.org/>

Tabulka 3.3: Vybrané generátory pro galerie

<b>Program</b>	<b>Jazyk</b>
Exposé <sup>16</sup>	shell
fgallery <sup>17</sup>	Perl
Prosopopee <sup>18</sup>	Python
Sigal <sup>19</sup>	Python

---

<sup>16</sup><https://github.com/Jack000/Expose/>

<sup>17</sup><https://gitlab.com/wavexx/fgallery/>

<sup>18</sup><https://github.com/Psycojoker/prosopopee/>

<sup>19</sup><http://sigal.saimon.org/>



## 4. Značkovací jazyky

Tato kapitola se zabývá analýzou nejběžnějších jazyků pro popis obsahu používaných ve statických generátorech z předchozí kapitoly, dalším důležitým jazykům ze světa sázení a jejich principům.

### 4.1 Principy značkovacích jazyků

Definici konceptu značkovacích jazyků nebo-li „markup jazyků“, můžeme najít například v RFC 7764<sup>1</sup>, tedy, že v počítačových systémech jsou kontextuální data ukládána a zpracovávána několika technikami. Informaci lze kódovat jako čistý text bez speciálních formátovacích znaků. Tento přístup je jednoduchý pro implementaci i použití, ovšem neumožňuje složitější formátování textu.

Kódovat můžeme i do binárních formátů určených ke zpracování a interpretaci specializovanými programy, jimiž mohou být například LibreOffice Writer, či proprietární software<sup>2</sup> Microsoft Word. Zřejmou nevýhodou je to, že zdroj není čitelný bez programu určeného pro jeho interpretaci.

Markup jazyky se snaží o spojení nejlepšího z obou světů, tedy o obsah s možností formátování, který je jednoduše čitelný jak pro člověka, tak pro stroj. Toho je dosaženo tím, že v je v běžných textových souborech přiřazen vybraným znakům speciální význam. Uživatel je schopen tyto znaky psát bez potřeby speciálních nástrojů, a tím jednoduše tento speciální význam vyjádřit. Například v rámci jazyka Markdown, jenž je popsán v následujících sekcích, se význam znaku # na začátku řádku mění z běžného křížku na definování nadpisu první úrovně. (Leonard, 2016)

---

<sup>1</sup>Jako *RFC* se označují standardy vydané organizací IETF (Internet Engineering Task Force).

<sup>2</sup>Software nerespektující základní svobody uživatele, které jsou popsány na adrese <https://www.gnu.org/philosophy/free-sw.html>

## 4.2 Nejběžnější jazyky

V současnosti existuje nespočet značkovacích jazyků, z nichž nejčastěji používaným je nepochybně HTML, ovšem tato práce se věnuje těm nejpoužívanějším jazykům, jež mají uživateli usnadnit psaní a sázení obsahu. Uživatel se tedy nemusí při jeho tvorbě nutně zabývat typografií a formátováním obsahu, což jsou aspekty, o které se později stará generátor pomocí šablon. U HTML je tomu naopak, uživatel řeší samotný obsah i formátování v jednu chvíli skrze různé druhy formátovacích tagů. O vyplňování obsahu do HTML se v případě staticky generovaných webových stránek stará právě samotný generátor.

Vybrané jazyky jsou zároveň cílené na čitelnost samotného zdrojového obsahu v čistém textu bez nutnosti jeho interpretace speciálním prostředím či zpracováním do jiných formátů, jimiž mohou být PDF, DjVu, PostScript apod. Například podtržení textu je v nějakém pseudo-jazyce reprezentováno opravdovým podtržením pomocí rovnítek, viz příklad 4.1, nikoliv obalením nadpisu ve speciální deklaraci, jako je tomu právě u jazyka HTML. Podtržení je poté pro čtenáře mnohem jasnější, jelikož nemusí přemýšlet, co v kontextu HTML daný tag znamená, kdežto podtržení z kontextu souboru vyplývá.

Fig. 4.1: Příklad podtržení nadpisu ve značkovacím pseudo-jazyce

```
Nadpis  
=====
```

Následující jazyky byly vybrány základě jejich výskytu u generátorů vybraných v předchozí kapitole a také podle aktuálních statistik ze serveru Slant (2020), který se věnuje obecnému určení oblíbenosti na základě hodnocení ze strany uživatelů.

### 4.2.1 $\text{\TeX}$

Tento jazyk se vzdaluje od původního konceptu čitelnosti zdroje, ovšem ve statických generátorech ho lze stále efektivně využít. Je jedním z nejrozšířenějších sázecích jazyků se spoustou možností a funkcionalit, z nichž velmi zajímavým rozšířením je prostředí pro psaní matematických formulí, díky němuž se jazyk stal velmi populárním v oblasti technických publikací. Tyto funkcionality se často objevují i v jiných jazycích a jsou

efektivně využívány pro jejich rozšíření.

Většina uživatelů se setkala spíše s jazykem  $\LaTeX$ , tedy s nadstavbou původního  $\TeX$ , která má uživateli zjednodušit práci svými makry a rozšířeními. Realita je ovšem taková, že  $\LaTeX$  dělá celou práci složitější, jak popisuje RNDr. Olšák:

Představte si, že si nějaký uživatel přečte  $\LaTeX$ ovou příručku a nabude dojmu, že mu bude stačit rozumět problematice sazby na úrovni této příručky. Pak se jednou překlepne třeba při sestavování tabulky a na terminálu na něj  $\TeX$  křičí: `Extra alignment tab has been changed to "\cr"`. Uživatel začne znovu listovat ve své příručce a zjistí, že tam o žádném `"\cr"` není jediná zmínka. Má pak tři možnosti: (1) Zmáčkne Enter a podobně se zachová i u dalších chyb. Pomyslí si, že ten  $\LaTeX$  je něco tajemného a mystického. (2) Propadne zoufalství a jde od toho. Dojde k závěru, že je lepší zůstat u Wordu. Vždyť stačí vzít tabulku v Excelu a jednoduše ji přemístit do Wordu a jaképak smolení se s nějakým podezřelým `"\cr"`. (3) Pořídí si  $\TeX$ book a po intenzivním studiu nakonec řekne: „aha“. V tuto chvíli ale už nepotřebuje, aby mu  $\LaTeX$  zakrýval složitost  $\TeX$ . (Olšák, 1997)

Ve výsledku je tedy lepší, z různých důvodů popsaných RNDr. Olšákem v jeho publikaci, použít samotný plain  $\TeX$  na úkor vyšší vstupní úrovně pro používání jazyka.

### 4.2.2 Troff

Troff je jedním z nejstarších jazyků a předchůdcem jazyka  $\TeX$ . Autorem původní verze je Joe Ossanna, po jehož smrti převzal vývoj Brian Kernighan. Samotný Troff je reimplementací a rozšířením původního programu RUNOFF z operačního systému CTSS. Vznikl za účelem sazby dokumentů na novém operačním systému Unix. (Ossanna – Kernighan, 1977)

Dnes existuje celá řada různých implementací a modernizovaných rozšíření, například Groff<sup>3</sup>, Heirloom troff<sup>4</sup>, nebo moderní Neatroff<sup>5</sup>, který se snaží o spojení toho nejlepšího

---

<sup>3</sup><https://www.gnu.org/software/groff/>

<sup>4</sup><http://heirloom.sourceforge.net/doctools.html>

<sup>5</sup><https://repo.or.cz/neatroff.git/>

ze všech předchozích implementací. Sám Brian Kernighan v soukromé emailové konverzaci doporučuje použití některé alternativní implementace, jež jsou podle Keringhana lepší ve všech směrech.

I přes vznik mnoha alternativních jazyků, například dříve zmíněného T<sub>E</sub>Xu a L<sup>A</sup>T<sub>E</sub>Xu, je Troff (Groff) stále hojně využíván v praxi, zejména u softwarové dokumentace v Unixových operačních systémech.

### 4.2.3 Markdown

Jazyk Markdown vznikl 19. března roku 2004, když John Gruber vydal první popis syntaxe a referenční implementaci.

Hlavním z cílů syntaxe jazyka je vytvářet co možná nejčitelnější obsah podobě čistého textu. Dokument psaný v Markdownu by měl být publikovatelný jako textový soubor bez dalších úprav a zpracování. Jazyk byl ovlivněn několika již existujícími specifikacemi jiných jazyků, ovšem největším zdrojem inspirace pro jeho vznik jsou čisté e-mailové korespondence<sup>6</sup>. (Gruber, 2004)

První specifikaci Gruber vydal společně s referenční implementací v jazyce Perl, která prováděla konverzi zdrojového souboru popsaného v jazyce Markdown do HTML. Tento program je také pojmenován *Markdown*, ovšem mluvíme-li o *Markdownu*, máme nejčastěji na mysli samotnou syntaxi. Ta má dnes mnoho implementací v různých programovacích jazycích. Gruberova specifikace ovšem není formálním standardem, kvůli čemuž vznikl veliký počet alternativních a více či méně pozměněných implementací, jež nemusí být navzájem kompatibilní. Nejčastějšími z nich jsou například Github Markdown, CommonMark, RMarkdown a mnoho dalších. (MacFarlane, 2019)

Nevyužívanější formální specifikací je právě CommonMark<sup>7</sup>, jenž slouží jako pevný základ většiny rozšíření. (Martí, 2017)

Podobně, jako je tomu u specifikací, existuje velké množství programů, které tyto různé specifikace překládají. Švýcarským nožem mezi nimi je program Pandoc<sup>8</sup>, jenž umí

---

<sup>6</sup>E-mail, který není formátován pomocí HTML.

<sup>7</sup><https://commonmark.org/>

<sup>8</sup><https://pandoc.org/>

překládat Markdown do enormního výběru jiných formátů, nebo z jiných formátů zpět. Tato funkcionality se nevztahuje pouze na jazyk Markdown, Pandoc dokáže operovat mezi všemi podporovanými formáty, například dokáže konvertovat obsah z HTML do T<sub>E</sub>Xu. Na druhou stranu existují i velmi jednoduché překladače, například program smu<sup>9</sup>, který umí překládat Markdown do HTML nebo čistého textu a neobsahuje více než šest set řádků kódu hlavního programu.

Užitečným rozšířením je, mimo jiné, také integrace matematického prostředí z jazyka T<sub>E</sub>X, viz sekce 4.2.1.

#### 4.2.4 Org-mode

Org-mode vznikl jako jeden z módů pro editor Emacs<sup>10</sup>. Funguje podobně jako ostatní mark-up jazyky, tedy jako jeden centrální systém pro správu obsahu, ze kterého lze vytvářet jiné formáty, například HTML, L<sup>A</sup>T<sub>E</sub>X, Open Document, Markdown, PDF a podobně s možností přidání libovolného nového backendu. Cílem Org-mode je možnost ho používat i s minimální úrovní jeho znalosti, ovšem jeho funkcionality je vždy přístupná. Vše je realizováno pouze na čistých textových souborech, nejlépe přenositelným typem souboru. Editor Emacs je zároveň velmi často portován na různé druhy systémů, a je tedy možné ho využívat v podstatě kdekoliv. (The Org Mode Developers, 2020)

Podporuje také „literate programming“ a „reproducible research“, tedy, že Org soubory mohou obsahovat plně funkční bloky s kódem, které lze hodnotit v rámci systému, a výstup bloků lze automaticky vkládat přímo do dokumentu. (Schulte et al., 2012)

Jak popisuje Dominik (2008) ve svém krátkém technickém popisu, Org-mode umí navrhování, psaní poznámek, hypertextové odkazy, tabulky, seznamy, plánování projektů, GTD, HTML a L<sup>A</sup>T<sub>E</sub>X, a to všechno v čistých textových souborech v editoru Emacs.

---

<sup>9</sup><https://github.com/Gottox/smu/>

<sup>10</sup><https://www.gnu.org/software/emacs/>

## 4.2.5 reStructuredText

Tento jazyk, známý také jako ReST, je, stejně jako Markdown, zároveň syntaxí i parsovacím systémem syntaxe pro tvorbu dokumentů a webových stránek. Svou oblibu získal hlavně v komunitě jazyka Python. Ve své dokumentaci<sup>11</sup> je popisován jako syntax pro využití ke psaní *Python docstrings* a dalších druhů dokumentace, která je spolehlivá a jednoduchá. ReST vznikl v návaznosti na jazyk StructuredText, jenž trpěl mnoha nedostatky. Cílem jazyka reStructuredText je tyto nedostatky opravit a doplnit. (Goodger, 2012)

S jazykem se lze setkat u značné části existujících generátorů statických webových stránek, z nichž některé jsou zmíněny v kapitole 3.

---

<sup>11</sup><https://docutils.sourceforge.io/rst.html>

# 5. Taxonomie požadavků pro modelový web

Tato kapitola se věnuje určení základních požadavků pro modelovou implementaci. Jsou zde shrnuta obecná kritéria, která platí pro většinu webových prezentací, a také kritéria specifická pro modelovou implementaci v rámci této práce. Dle těchto kritérií je poté v následující kapitole vytvořena samotná implementace.

Jako modelová implementace byl zvolen web pro distribuci výukových materiálů a odkazů užitečných pro výuku. Tvorba těchto webových stránek je zadána Ústavem výzkumu a rozvoje vzdělávání Pedagogické fakulty Univerzity Karlovy za účelem usnadnění práce již aktivních učitelů v době šíření viru COVID-19. Tyto webové stránky mají učitelům pomoci s přípravou distanční výuky a úkolů v době vyhlášení stavu nouze a celostátní karantény. Modelová implementace je tedy plně využívána v praxi mnoha pedagogy z celé republiky. Tuto implementaci lze ovšem použít pro distribuci jakýchkoliv jiných výukových materiálů či ke psaní a správě dokumentace.

## 5.1 Obecná kritéria

Jako zdroj obecných kritérií je použit článek ze serveru Calomel (2017), který se mimo jiné věnuje i optimalizacím, jež jsou dále popsány v sekci 6.5.

Z důvodu potenciálního vytížení sítě je nutné, aby byl celý obsah optimalizován za účelem předejití vysoké latence, a to z důvodů probíraných v předchozí části práce, tedy v sekci 2.1.

Stránky by měly být udržovatelné i po předání jinému správci, a celý systém by tedy měl být dostatečně zdokumentován. Také je důležité, aby byla zajištěna kompatibilita s nejběžněji používanými prohlížeči.

## 5.2 Kritéria specifická pro modelový web

Specifická kritéria jsou vytvořena na základě požadavků autorů obsahu, tedy učitelů, z nichž má každý své specifické požadavky na funkce a vlastnosti, které musí obsah splňovat. Následující kritéria jsou souhrnem a kompromisem mezi všemi požadavky.

Stránky musí být staticky generované a není tedy žádoucí v rámci webu řešit uživatelské účty, přihlašování apod. Hlavním požadavkem pro strukturu stránky je možnost dělit obsah na sekce dle druhu školy (základní škola, střední škola, vysoká škola atd.) a dále pak na subsekce podle předmětů a oborů.

Do samotného obsahu musí být možné vkládat přílohy ke stažení v různých formátech, obrázky a videa s možností jejich ocitování, tedy uvedení autora, názvu díla apod. Všechny přiložené soubory musí být distribuovatelné přímo z webových stránek, nikoliv z externích zdrojů. Všechna videa je nutné vložit do stránky a musí je být možné přehrát v nativním přehrávači prohlížeče bez nutnosti otevírání externích webových stránek či programů. V hlavičce každé stránky musí být možné specifikovat následující informace: autora či seznam autorů obsahu, skupinu, pro kterou je obsah určen a časovou dotaci.

Obsah stránek musí být možné spravovat předem pověřenými uživateli a jeho změny musí být zaznamenávány v decentralizovaném verzovacím systému. Generování statického webu na základě změn obsahu je nutné řešit automatizovaně — bez dalších zásahů správce či manuálního nahrání nového obsahu na webserver.

## 5.3 Kritéria pro šablony a design

Obsah musí být zobrazitelný na každém druhu zařízení, tedy jak na monitorech s nadstandardní velikostí, tak na mobilních zařízeních. Zároveň musí být snadno čitelný, v nejlepším případě vysoce kontrastní černý text na bílém pozadí s dostatečnou velikostí. Navigace v obsahu musí být jednoduchá a intuitivní a vzhled celé stránky konzistentní. Na stránce nesmí přesahovat objem vizuálních elementů nad obsahem. Relevantní obsah by měl být na jednom místě, nikoliv rozdělený na několik různých stránek, mezi kterými musí uživatel přecházet.



## 6. Modelová implementace

Tato část práce se věnuje tvorbě modelové implementace systému pro generování statických webových stránek dle definovaných požadavků v kapitole 5. Jsou zde vybrány vhodné součásti, ze kterých je modelová implementace složena.

### 6.1 Výběr vhodného systému

Modelový systém se skládá ze dvou částí, a to z verzovacího systému pro správu obsahu a generátoru statického HTML. Obě tyto součásti jsou vybírány na základě poznatků z předchozích částí práce.

#### 6.1.1 Verzovací systém pro správu obsahu

Pro správu obsahu i šablon a statických souborů byl zvolen distribuovaný verzovací systém Git, jenž má v porovnání s jinými verzovacími systémy, zejména centralizovanými, spousty výhod. Jeho hlavní výhodou je rozšířené využití v praxi a snadné používání. Díky svým decentralizovaným vlastnostem ho lze využívat v mnoha odlišných pracovních postupech. S naklonovaným repozitářem lze pracovat i bez připojení k síti, což lze považovat i za druh zálohy. Git také umožňuje slučování různých změn od mnoha uživatelů a dovoluje jednoduše řešit potenciální konflikty. (Chacon, 2009)

Skvěle využitelnou funkcí pro modelovou implementaci je také to, že po provedení změn v repozitáři lze pomocí Gitu spouštět skripty, které mohou provádět automatické generování obsahu a další užitečné operace. Tato funkcionalita je implementována v rámci modelové implementace v sekci 6.6.1.

#### 6.1.2 Generátor statického webu

Protože forma modelového webu odpovídá paradigmatu webové prezentace ze sekce 3.1, byl pro jeho generování použit program Zola<sup>1</sup>.

---

<sup>1</sup><https://www.getzola.org/>

Vybraný generátor splňuje všechny požadavky z kapitoly 5 a oproti jiným systémům je výhodný tím, že je napsaný v jazyce Rust. Je tedy mnohem rychlejší a bezpečnější než většina jeho alternativ (Gouy, 2020). Tato výhoda v rychlosti se znatelně projevuje při zpracování obsáhlých webových stránek či mnoha obrázků.

V rámci generátoru je využíván značkovací jazyk Markdown, který je snadný a velmi rozšířený. Kromě těchto výhod si zachovává většinu funkcí a rysů, jež lze najít v ostatních složitých systémech. Zároveň je možné generátor zkompilevat<sup>2</sup> do jednoho staticky linkovaného<sup>3</sup> binárního souboru, s nímž se pracuje mnohem lépe než se složitým frameworkem.

## 6.2 Tvorba šablony

Jak se uvádí v dokumentaci<sup>4</sup>, Zola pracuje s několika druhy stránek, primárně s takzvanou *sekcí* a *stránkou*.

*Stránka* slouží pouze k předání obsahu a nikoliv k dalšímu větvení struktury. Dá se tedy říci, že stránka v rámci stromové struktury reprezentuje konec větve.

Každá *sekce* může mít vlastní obsah, ovšem může obsahovat i další subsekce, pomocí nichž lze obsah ve stromové struktuře větvit. Kořenem celého stromu je speciální sekce s názvem *index*.

Pro každou část se obvykle používá vlastní HTML šablona, ovšem není to pravidlem a každá část větve může využívat šablonu jinou. To je užitečné například u stránek s různými druhy obsahu. V rámci modelového webu zůstává druh obsahu stejný a není tedy třeba se odchylovat od standardní struktury.

Soubory se šablonami se nachází ve složce `templates/`, ve které generátor vždy očekává šablonu `index.html`. Ta se využívá jak k vykreslení úvodní kořenové stránky, tak jako základ, jež mohou ostatní šablony rozšiřovat. Tato kořenová šablona tedy obsahuje základní strukturu celé stránky, přičemž navazující šablony jen mění určité její části a nedefinují celou strukturu znovu.

---

<sup>2</sup>Spustitelné programy musí být převedeny do strojového kódu procesem zvaným *kompilace*.

<sup>3</sup>Spustitelný soubor má všechny potřebné knihovny integrované v sobě.

<sup>4</sup><https://www.getzola.org/documentation/content/overview/>

Generátor Zola v šablonách hledá vlastní řídicí sekvence, jež se popisují kombinací složených závorek a dalších znaků. Existují tři druhy kombinací, které lze použít:

- `{% %}` – Metoda, funkce, cykly, podmínky, práce s proměnnou atd.
- `{{ }}` – Výpis do HTML
- `{# #}` – Komentář

Generátor také vyžaduje konfigurační soubor `config.toml` v kořenové složce projektu, jenž obsahuje různá nastavení stránky, globální proměnné a chování generátoru.

Fig. 6.1: Příklad jednoduché konfigurace v souboru `config.toml`

```
# Adresa pro kterou se generují odkazy
base_url = "https://ucitelonline.pedf.cuni.cz"
# Název stránky
title = "Učitel online"
# Popis stránky
description = "Web pro distribuci užitečných materiálů"
# Zda se bude zpracovávat CSS systémem Sass
compile_sass = true
```

Systém vždy zpracuje úvodní šablonu `index.html`, ze které pak lze odvíjet ostatní šablony. Tato hlavní šablona obsahuje strukturu celé webové stránky a nesmí v ní chybět validní HTML struktura, tedy hlavička, tělo, metadata, kódování a podobně. Do struktury lze vkládat libovolné řídicí sekvence generátoru, které ovlivňují výsledný výstup.

Fig. 6.2: Základní šablona `index.html`

```
<!DOCTYPE html>
<html lang="cs">
<head>
  <meta charset="UTF-8">
  <title>{{ config.title }}</title>
</head>
<body>
```

```
</body>
</html>
```

V příkladu 6.2 je název stránky mezi tagy `<title></title>` vyplněn generátorem. Ten do šablony vloží hodnotu konstanty `config.title`, která je nastavena v konfiguračním souboru `config.toml` z příkladu 6.1. Názvem stránky bude tedy řetězec „Učitel online“. Generátor dokáže převzít kteroukoliv konstantu z kontextu konfiguračního souboru.

Všechny tyto řídicí sekvence, nebo také *direktivy*, lze v rámci generátoru navazovat na sebe, podobně jako je tomu v unixových systémech. Spojování funkcí a filtrů se provádí znakem `|`, stejně jako v POSIX<sup>5</sup> shellu, kde se výstup jednoho příkazu stane vstupem příkazu navazujícího. Například je možné název stránky vypsat ve velkých písmenech i přesto, že v konfiguračním souboru je řetězec s názvem formátován pouze s velkým písmenem na začátku. K převedení na velká písmena slouží filtr `upper`. Názvem stránky bude po zpracování programem 6.3 řetězec „UČITEL ONLINE“.

Fig. 6.3: Základní šablona s filtrem pro přepsání názvu na velká písmena

```
<!DOCTYPE html>
<html lang="cs">
<head>
  <meta charset="UTF-8">
  <title>{{ config.title | upper }}</title>
</head>
<body>
</body>
</html>
```

Generátor umožňuje v šabloně vytvářet speciální bloky, jejichž obsah lze v navazujících šablonách měnit. Je tedy možné měnit části struktury šablony. K vysvětlení principu fungování bloků je název stránky v příkladu 6.4 obalen blokem `title` a do těla vložen blok `content`.

<sup>5</sup>Portable Operating System Interface – Rodina standardů unixových systémů

Fig. 6.4: Využití bloků v šabloně z příkladu 6.3

```
<!DOCTYPE html>
<html lang="cs">
<head>
  <meta charset="UTF-8">
  <title>{% block title %}{{ config.title | upper }}{%
    endblock %}</title>
</head>
<body>
{% block content %}
  Ahoj, světe!
{% endblock %}
</body>
</html>
```

Název stránky zůstane stejný a v jejím těle přibude text „Ahoj, světe!“. Vytvoříme-li novou šablonu s názvem `section.html`, generátor nám umožní rozšířit ji o původní šablonu `index.html` a měnit pouze definované bloky. Není tedy nutné znovu definovat celou strukturu stránky. Pro importování nebo-li rozšíření šablony slouží direktiva `extends`.

Fig. 6.5: Definice nové šablony `section.html` rozšiřující šablonu z příkladu 6.4

```
{% extends "index.html" %}
{% block title %}{{ config.title | upper }} &ndash; {{ section.
  title }}{% endblock %}
{% block content %}
  Toto je obsah kategorie.
{% endblock %}
```

Šablona `section.html` se v rámci generátoru Zola implicitně využívá pro všechny existující sekce<sup>6</sup>. Názvem stránky v této šabloně bude, podobně jako u hlavní šablony, název stránky z konstanty `config.title` definované v konfiguračním souboru, ale také spojovník a název dané sekce. Za modelový výstup lze považovat například „UČITEL

<sup>6</sup><https://www.getzola.org/documentation/content/section/>

ONLINE – základní a střední škola“, bude-li se uživatel nacházet v sekci pro základní a střední školy.

V bloku s obsahem bude původní obsah „Ahoj, světe!“ nahrazen za řetězec „Toto je obsah kategorie“. Ten ovšem nechceme definovat přímo v šabloně, nýbrž cílem generátoru je vyplňovat obsah ze zdrojových souborů v sázecím jazyce, viz sekce 2.2. Zola pro vkládání obsahu využívá stejný princip jako v ostatních případech, tedy vypsání obsahu proměnné, v tomto případě proměnné `section.content`, která obsahuje zkompilevané HTML z daného Markdown souboru. Zároveň je dobrou praktikou provést vyčištění vstupu filtrem `safe`<sup>7</sup>.

Fig. 6.6: Vkládání obsahu ze zdrojového Markdown souboru

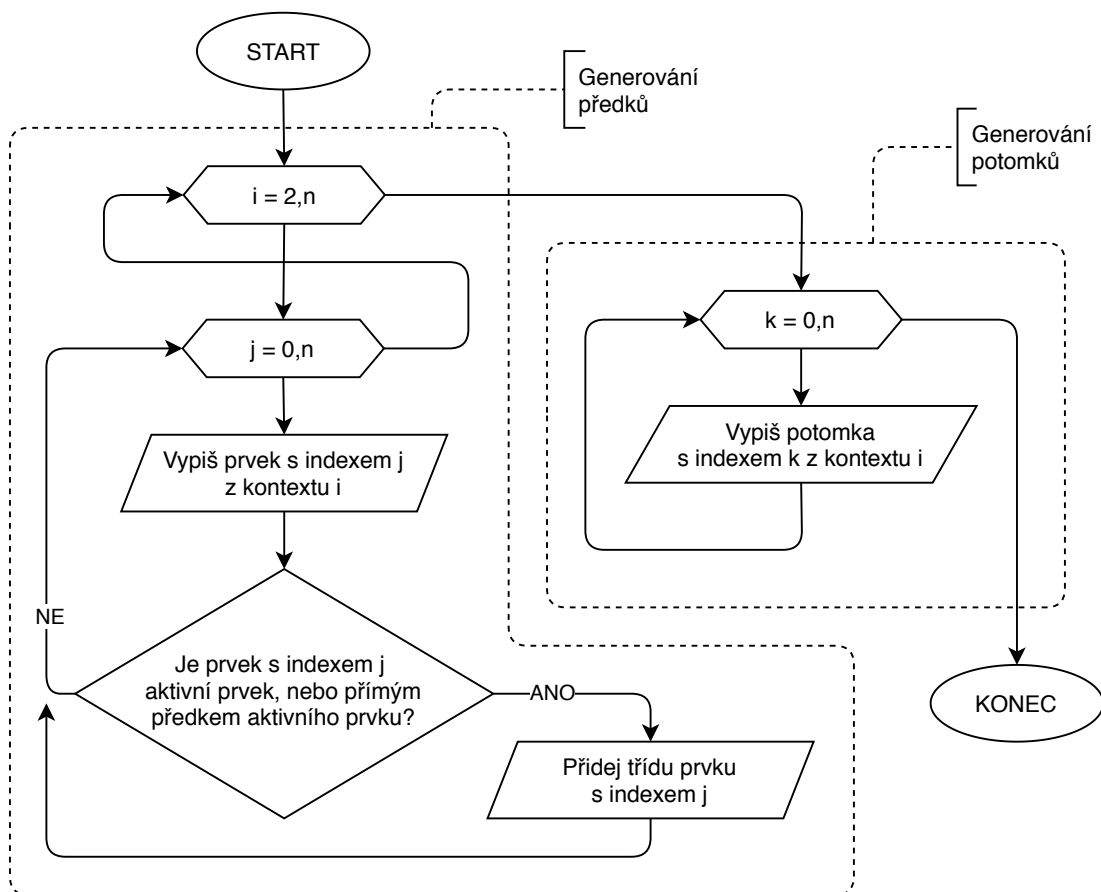
```
{% extends "index.html" %}
{% block title %}{{ config.title | upper }} &ndash; {{ section.
    title }}{% endblock %}
{% block content %}
    {{ section.content | safe }}
{% endblock %}
```

Z důvodu přenositelnosti systému a principu staticky generovaných stránek by žádný obsah neměl být definován přímo v šabloně, nýbrž by měl být do stránky vkládán generátorem z proměnných nebo ze sázeného obsahu. V rámci modelové implementace je toto nepsané pravidlo dodržováno.

## 6.3 Automatické generování vícevrstvé navigace

Obsah modelové implementace je dělen do stromové datové struktury o potenciálně nekonečné hloubce, kdy každá část větve je v rámci generátoru vlastní kategorií, nikoliv stránkou. Pro modelovou implementaci bylo zvoleno, aby se navigace generovala v návaznosti na aktivní cestu ve stromě. Ve stránce jsou dvě různé navigace: hlavní, která je vždy viditelná a obsahuje rozdělení obsahu dle škol, a vedlejší, jež zobrazuje aktivní větev stromu.

<sup>7</sup><https://tera.netlify.com/docs/#safe>



Obrázek 6.1: Diagram průběhu generování vícevrstvé navigace

První vrstvou struktury jsou hlavní sekce, v rámci implementace pojmenované jako  $L_1$ , které jsou vždy vypsány ve vlastní navigaci. Pod touto navigací je zobrazen seznam všech kategorií, které vybraná položka v  $L_1$  obsahuje. Pokud uživatel zvolí kteroukoliv položku v  $L_2$ , v navigaci se objeví další sloupec, který obsahuje všechny podkategorie vybrané položky, tedy všechny podkategorie ve vrstvě  $L_3$ . Takto lze stromem procházet potenciálně donekonečna. Styly modelové šablony ovšem počítají s maximální hloubkou čtyř subkategorií.

Tato funkcionální je implementována pomocí tří cyklů, z nichž jeden je vložený. První cyklus (příklad 6.7) se provádí pro všechny rodiče aktivní kategorie vrstev  $L_2, L_3, \dots, L_n$ , kde  $n$  je aktuální vrstva. V každé iteraci se mění kontext, ve kterém generátor pracuje. Z daného kontextu generátor vypisuje pomocí vnořeného cyklu všechny subkategorie. Ve druhém cyklu (příklad 6.8) se vypisují všichni potomci dané stránky, tedy potomci ve vrstvě  $L_{n+1}$ .

Fig. 6.7: Cyklus pro vypisování všech rodičů v dané větvi navigace

```
{% if section.ancestors %}
  {% for s in section.ancestors %}
    {% if loop.index < 2 %}{% continue %}{% endif %}
    <ul>
      {% set s = get_section(path=s) %}
      {% for s in s.subsections %}
        {% set s = get_section(path=s) %}
        <li><a href="{ s.permalink }"
          {% if current_path == s.path %}
            class="active"
          {% elif current_path is containing(s.path) %}
            class="ancestor"
          {% endif %}
        >{{ s.title }}</a></li>
      {% endfor %}
    </ul>
  {% endfor %}
{% endif %}
```

Fig. 6.8: Cyklus pro vypisování všech potomků dané stránky do navigace

```
{% if section.subsections %}
  <ul>
    {% for s in section.subsections %}
      {% set s = get_section(path=s) %}
      <li><a href="{ s.permalink }">{{ s.title }}</a></li>
    {% endfor %}
  </ul>
{% endif %}
```



## 6.4 Rozšíření šablony

Ve výchozím stavu neumí generátor zpracovávat nic jiného, než co je uvedeno ve specifikaci CommonMark, viz sekce 4.2.3. Dle požadavků modelového webu je nutné, aby generátor uměl vkládat videa přímo do stránky. Taková funkcionality není součástí specifikace CommonMark, a je tedy potřeba rozšířit generátor. Nejvhodnějším způsobem přidání vlastních funkcionalit je využití filtrů, jež se v rámci generátoru nazývají *shortcode*.

Principem vlastních filtrů je to, že si uživatel vytvoří vlastní šablonu, kterou lze vyvolat pomocí speciální řídicí sekvence přímo z obsahu. Každý tento shortcode může pracovat s libovolným množstvím proměnných a po zpracování vloží do místa vyvolání zkompileovaný HTML kód. Dá se tedy říci, že shortcode je ve své podstatě funkce, jež umí pracovat s parametry.

Pro tvorbu těchto filtrů je v generátoru Zola určena složka `templates/shortcodes` obsahující jejich HTML šablony a kód pro zpracování generátorem. Název HTML souboru definuje název vlastního filtru. Vytvoříme-li uvnitř této složky soubor nazvaný `video.html`, budeme v obsahu schopni využívat vlastní filtr s názvem `video`.

Fig. 6.9: Příklad jednoduchého filtru s jedním atributem

```
<video controls><source src="{{ src }}"></video>
```

V příkladu 6.9 bude filtr očekávat atribut `src` a bude vracet jednoduchý HTML kód pro vložení videa do stránky. Tento filtr lze vyvolat kdekoliv v obsahu, tedy v kterémkoliv souboru s koncovkou `.md`, v nichž je uložen obsah. Za názvem filtru se do závorky uvádí parametry. Pro lepší přehlednost lze parametry oddělovat čárkou, což ovšem není pro správné zpracování generátorem nutné.

Fig. 6.10: Vyvolání vlastního filtru s jedním parametrem

```
{{ video(src="video.webm") }}
```

V rámci vybraného generátoru není nutné specifikovat atributy na jeden řádek a lze je pro další zpřehlednění vypisovat na více řádků, stejně jako v příkladu 6.13. Výstupem této direktivy bude následující HTML kód.

Fig. 6.11: Výstup direktivy z příkladu 6.9

```
<video controls><source src="video.webm"></video>
```

Součástí požadavků pro modelový web jsou i citace přiložených souborů a videí. Existující filtr je tedy třeba rozšířit o možnost přiložení různých metadat. Tato metadata ovšem nejsou pro vložení videa povinná. Ve specifikaci vlastních filtrů lze využívat všechny operátory, jež generátor nabízí. Nejlepším přístupem k tomuto problému je tedy využití jednoduchých podmínek, které kontrolují, zda je každá z hodnot zadána jako parametr, a v případě že ano, vepíše se do obsahu. Atributy ošetřené podmínkami tedy nejsou povinné, zatímco nevyplněný atribut `src` by při generování vyvolal chybu. V následujícím příkladu jsou přidány podmínky pro kontrolu a případné vložení, jimiž jsou název videa (`title`), jméno autora (`author`) a rok vytvoření (`year`).

Fig. 6.12: Filtr pro vkládání videa s využitím podmínek

```
<video controls><source src="{{ src }}"></video>
{% if title or year and author %}
<div class="metadata">
    {% if title %}{{ title }}{% endif %}
    {% if author and year %}
        ({{ year }}, {{ author }})
    {% endif %}
</div>
{% endif %}
```

Filtr je opět možné vyvolat pomocí stejné direktivy kdekoliv v obsahu, ovšem nyní lze libovolně přidávat parametry pro metadata.

Fig. 6.13: Vyvolání filtru 6.12 s formátováním na řádky

```
{{ video(
    src="video.webm"
    title="Název videa"
    author="Jméno autora"
    year="2020"
) }}
```

Protože byly zadány všechny povinné i nepovinné atributy, výstupem toho filtru budou i části kódu s metadaty.

Fig. 6.14: Výstup direktivy z příkladu 6.13

```
<video controls><source src="video.webm"></video>
<div class="metadata">
  Název videa (2020, Jméno autora)
</div>
```

Pro modelový web byla zvážena možnost vypisování obsahu automaticky, tedy že program projde složku s obsahem a pokud narazí na soubor se specifikovanou koncovkou, vypíše jej do obsahu podle daných pravidel. Generátor Zola umožňuje prohledávání složek a práci se soubory, pro které se v rámci Zoly používá termín „asset“. Tuto funkcionalitu lze tedy implementovat jednoduchým cyklem a filtrem, které zpracují všechny případné soubory ve složce dané stránky. Soubory lze filtrovat mnoha způsoby, z nichž je nejuniverzálnější funkce `matching()`, která dovoluje filtrovat vstup regulárními výrazy dle jejich implementace v jazyce Rust<sup>8</sup>. V následujícím příkladu je pro ilustraci této funkcionality implementován program vypisující obrázky s předem definovanými koncovkami.

Fig. 6.15: Automatický výpis obrázků s pevně definovanými koncovkami

```
{% if section.assets %}
  {% for asset in section.assets %}
    {% if asset is matching("\.(?i:jpg|gif|png)$") %}
      
    {% endif %}
  {% endfor %}
{% endif %}
```

Toto řešení ovšem není ve výsledném modelu implementováno, protože jedním z požadavků je možnost vkládání souborů na libovolné místo v obsahu. Na stejném principu je vytvořen filtr pro vkládání souborů, který tento požadavek splňuje. Výhodou filtru je,

<sup>8</sup><https://docs.rs/regex/1.3.6/regex/>

že ho lze vyvolat kdekoliv v obsahu a není vázán na pevně dané místo v šabloně. Filtr očekává alespoň jeden parametr uvádějící název souboru bez koncovky, podle něhož pak vyhledá všechny různé formáty s tímto názvem, a ty vloží do stránky. Druhým libovolným parametrem je název souboru, který se do stránky vloží místo názvu souboru. To umožňuje uživateli volně pracovat s názvy souborů v souborové struktuře bez ovlivnění obsahu stránky.

Fig. 6.16: Filtr pro výpis souborů s automatickým hledáním

```
{% if section.assets and filename %}
<div class="file">
  <div class="title">
    {% if title %}
      {{ title }}
    {% else %}
      {{ filename }}
    {% endif %}
  </div>
  {% for asset in section.assets %}
    {% if asset is matching(section.path ~ filename ~ "\..*"
      "$") %}
      <a href="{{ get_url(path=asset) }}" class="format">
        {{ asset | split(pat=".") | last }}</a>
      {% endif %}
    {% endfor %}
  </div>
{% endif %}
```

V první části filtr zkontroluje, zda byl vyplněn parametr `title`: pokud ano, nastaví ho jako název souboru v obsahu, v opačném případě využije název souboru samotného.

Ve druhém kroku nastává kontrola, jestli se ve složce nacházejí soubory (mimo hlavní soubor `_index.md`) — když ano, tak se iterativně zkontrolují všechny soubory, zda splňují podmínku názvu. Kontrola této podmínky je tvořena kombinací proměnných generátoru a regulárního výrazu. Každý soubor, který splňuje podmínku, je poté vyp-

sán do obsahu jako přímý odkaz k jeho stažení.

Jako text v odkazu se použije koncovka souboru, která se získává spojením několika filtrů, tedy filtru `split(pat=".")`, který rozdělí řetězec podle znaku tečka do pole, a navazujícího filtru `last`, jenž vrátí poslední položku v poli. Tím filtr získá samotnou koncovku souboru.

Filtr lze vyvolat stejně, jako je tomu u filtru pro vkládání videa. Název filtru je opět definován názvem souboru `templates/shortcodes/document.html` a bude jím tedy název `document()`.

Fig. 6.17: Vyvolání filtru 6.16

```
{{ document(
    filename="pracovni-list"
    title="Pracovní list"
) }}
```

V příkladu 6.17 je definován i nepovinný atribut `title`, který pro přehlednost umožňuje nastavit název. Atribut `filename` definuje název souboru ve složce bez koncovky. Všechny soubory, jež chce uživatel vypsat, musí tedy mít stejný název a musí se lišit pouze koncovkou. Jsou-li ve složce soubory s názvem `pracovni-list` a koncovkami `pdf`, `odt`, `djvu` a `ps`, bude výstupem filtru následující HTML.

Fig. 6.18: Výstup direktivy z příkladu 6.17

```
<div class="file">
    <div class="title">Pracovní list</div>
    <a href="pracovni-list.pdf">pdf</a>
    <a href="pracovni-list.odt">odt</a>
    <a href="pracovni-list.djvu">djvu</a>
    <a href="pracovni-list.ps">ps</a>
</div>
```

## 6.5 Optimalizace

Optimalizace je provedena na základě článku ze serveru Calomel (2017), který se věnuje sestavením užitečných rad pro optimalizaci webových stránek na serverech s omezeným připojením do sítě a zvýšení spokojenosti uživatelů z užívání optimalizovaného webu, jak je rozebráno v sekci 2.1.

Jak se na webu Calomel píše, provozování webserveru může být hodnotnou zkušeností, ale zároveň může být i zkouškou trpělivosti. Chcete svým uživatelům předávat všechny vaše stránky a obrázky, ovšem máte jen omezenou šířku pásma, pomocí které můžete data přenášet. Pokud přetížíte své připojení, klienti navštěvující váš web server si budou myslet, že je pomalý a neresponzivní. Je tedy třeba webový server nastavit tím nejlepším možným způsobem s cílem získat co nejvíce návštěv a zlepšit zážitek vašim návštěvníkům. Následující rady slouží ke snížení zátěže serveru, ke zrychlení odesílání stránek a k zastavení nechtěného a škodlivého provozu.

Práce se věnuje pouze technickým optimalizacím spojeným s tvorbou samotné webové stránky, nikoliv však optimalizacím sítě, web serveru a vizuálního návrhu. Nenačítá-li se stránka během několika vteřin, většina uživatelů jednoduše odejde. Cílem této sekce je provést optimalizace, které urychlí načítání modelové implementace.

### 6.5.1 Typy a kvalita obrázků

Fotografie a grafika využívají mnohem více dat pro přenos než běžný HTML text a je tedy nutné provést optimalizaci (kompresi) obrázků na co nejmenší možnou velikost souborů. Obrázky není třeba renderovat na více než 72 dpi a pro každý druh grafiky je třeba zvolit vhodný formát, tj. formát JPEG pro fotografie a formáty PNG či SVG pro jednoduchou grafiku. Rastrové obrázky by neměly přesáhnout maximální rozlišení zobrazované na stránce. Klíčové je také nevyužívat obrázky v případě, kdy je lze nahradit čistým HTML a CSS.

Obrázky ve formátu JPEG mají velice efektivní ztrátovou kompresi, pomocí které lze zredukovat velikost obrázku o značnou část. Autor článku tvrdí, že většinu obrázků lze komprimovat až o 50% bez viditelné ztráty na kvalitě. Svě obrázky dokonce zkompri-

moval ze 27 kilobajtů na pouhých 8 kilobajtů s JPEG kompresí 60%.

### 6.5.2 Ikona *favicon.ico*

Původně je *favicon.ico* výtvořem firmy Microsoft, kdy Internet Explorer automaticky odesílal požadavek na pevnou URL `/favicon.ico` od kořene webového serveru. Jde o malou ikonku, jež se dnes zobrazuje u každé záložky s webovou stránkou. Problémem je, že se požadavkům na ní nelze vyhnout a vždy se počítá s tím, že ikona webu existuje. Odesílá se vždy s každou stránkou a některé prohlížeče se po ní dotazují z neznámých důvodů dvakrát. Autor článku uvádí, že u některých serverů bylo až 30% přenesených dat využito jen na odesílání ikony.

Principem optimalizace je udržet ikonu co nejmenší, v nejlepším případě tak malou, aby se vešla do jednoho TCP paketu, tedy do velikosti 1460 bajtů na většině systémů. Toho lze docílit tím, že ikona nebude větší než 16x16 pixelů s nízkou barevnou hloubkou, nejlépe pouze se čtyřmi barvami. Také je možné poslat jen 1x1 pixelů veliký prázdný obrázek nebo vracet stavový kód 204<sup>9</sup> a neodesílat ikonu žádnou.

### 6.5.3 Obecné HTML optimalizace

Redukcí nepotřebných znaků v HTML lze také ušetřit značnou část přenesených dat. Nejvhodnější je:

- nepoužívání HTML komentářů,
- využití CSS pro formátování stránek,
- využití oddělovače nebo elementu `span` namísto tabulek,
- odstranění přebytečných tagů a prázdných mezer a řádků,
- vytvoření obrázkových náhledů namísto odesílání obrázků v plném rozlišení,
- recyklování již použitých obrázků a tlačítek.

---

<sup>9</sup>204 No Content – Server úspěšně zpracoval požadavek, ale nevrací žádný obsah.

K odstranění přebytečných mezer, zalomení řádků, HTML komentářů a prázdných řádků lze použít automatický filtr pro kompresi výstupu. Toto by se nabízelo jako jedna z dalších možností pro implementaci rozšíření. Generátor Zola provádí kompresi CSS, ovšem nemá zabudovanou funkcionalitu pro minifikaci výsledného HTML, která je v době psaní této práce vyvíjena<sup>10</sup>.

Touto redukcí lze ušetřit 2% přenosu dat oproti ručně psanému neoptimalizovanému kódu. Je-li průměrná velikost stránky sto kilobajtů, lze touto optimalizací ušetřit dva kilobajty při každém odeslání stránky. Při odeslání sta tisíce stránek za měsíc je ve výsledku ušetřeno dvě stě megabajtů dat, které jsou jinak zbytečně odesílány uživatelům, kteří je stejně nezobrazí.

Další obecné rady pro optimalizaci HTML jsou uvedeny na serveru Yahoo! (2020), kde se uvádí spousta dalších způsobů ke zrychlení načítání stránky a k nižšímu vytížení sítě.

Připojením externích CSS a JavaScript souborů je umožněno jejich ukládání do paměti cache, což snižuje HTTP požadavky vůči serveru. Je-li obsah těchto souborů přímo ve stránce, je odeslán pokaždé s novou stránkou, a to vede ke zbytečnému vytěžování sítě. S tím souvisí i velikost stránek, kdy se soubory s větší než danou maximální velikostí do mezipaměti neukládají, a je proto dobré tuto velikost nepřekračovat.

Připojením externího CSS přímo do hlavičky je umožněno progresivní vykreslování webové stránky, které urychluje „Time To First Byte“, viz sekce 2.1. Naopak umístěním případných JavaScript souborů až na konec celé stránky se prioritizuje načítání viditelného obsahu před méně důležitými skripty.

#### 6.5.4 Optimalizace videa

Protože v modelové implementaci jsou do stránky vkládána i videa, je nutné provádět jejich optimalizaci podobně, jako je tomu u obrázků. Důležité je používat kvalitní kompresi, pouze nutné rozlišení a renderovat videa ve správném poměru stran a bez zbytečných černých okrajů. Při zpracování videa je dobré neprovádět jeho transkódování do jiného formátu z původního, ovšem zároveň je třeba dbát na kompatibilitu

---

<sup>10</sup><https://github.com/getzola/zola/issues/542/>



s prohlížeči, které různé formáty a kontejnery neumí vždy nativně přehrát.

## 6.6 Správa obsahu a verzování

Statické stránky neumožňují správu uživatelů v rámci webové aplikace, tedy to, že se případný editor nebo administrátor přihlásí a upravuje obsah klikáním či psaním v editoru, který během psaní formátuje text tak, jak bude ve výsledku vypadat. Správu uživatelů lze jednoduše řešit omezením přístupu na web server, kde mohou do obsahu zasahovat jen oprávnění uživatelé. To je však velmi těžkopádné řešení, protože neumožňuje práci více uživatelů najednou a neudrжуje předešlé verze obsahu a historii úprav. Lepší alternativou je využití některého verzovacího systému. Pro účely modelové implementace byl vybrán distribuovaný verzovací systém Git, jak je vysvětleno v sekci 6.1.1.

V tomto systému jsou soubory uloženy v repozitářích, kde každý projekt je vlastní repozitář. V rámci jednotlivých repozitářů se ukládají všechny změny obsahu prostřednictvím takzvaných *commitů* — záznamů o provedených změnách včetně jejich krátkého popisu a autora. Tyto revize lze provádět v různých větvích repozitáře a větve je možné mezi sebou spojovat a kombinovat. Rovněž je možné se vracet do kteréhokoliv bodu v historii v rámci každé větve.

Nastane-li konflikt při nahrávání změn, umožňuje Git jejich snadné vyřešení. Konflikt je stav, kdy například dva různí uživatelé provedli úpravy na stejném místě stejného souboru a snaží se je nahrát do repozitáře. Git v tuto chvíli druhého uživatele upozorní, že původní soubor byl změněn a je třeba tento konflikt vyřešit. Zamezuje se tak přepsání změn prvního uživatele.

K systému Git existují různé služby, jež tento systém rozšiřují o webové grafické rozhraní s množstvím dalších rozšíření. Nejčastěji používanými službami jsou GitHub<sup>11</sup>, GitLab<sup>12</sup>, nebo Bitbucket<sup>13</sup>, z nichž některé lze provozovat na vlastním serveru. Snad-

---

<sup>11</sup><https://github.com/>

<sup>12</sup><https://gitlab.com/>

<sup>13</sup><https://bitbucket.org/>

ným systémem pro vlastní provozování je také program Gitea<sup>14</sup>, jenž je oproti předem zmíněným systémům zcela svobodným softwarem<sup>15</sup> a je velmi jednoduchý na instalaci a správu. Tyto systémy mají navíc integrovaný jednoduchý editor pro úpravu souborů přímo z webového rozhraní a také umí zobrazovat náhled souborů s obsahem napsaným v jazyce Markdown, jenž je popsán v sekci 4.2.3.

### 6.6.1 Automatizace generování obsahu

Tato část práce se věnuje samotné implementaci automatického generování obsahu na základě změn v repozitáři.

Jak bylo zmíněno v sekci 6.1.1, git umožňuje nastavení takzvaných *Git hooks*, jež se v určité chvíli spustí. Jak uvádí dokumentace<sup>16</sup>, existuje spousta druhů hooků, které jsou vyvolány v různé části zpracování požadavku. V případě této implementace je nejvhodnější hook *post-receive*, jenž je spouštěn až po nahrání a zpracování všech změn v repozitáři.

Skript 6.19 po vyvolání Gitem provede veškeré potřebné operace ke zpracování nového obsahu na web serveru. Je složen z několika částí.

Jako první probíhá na řádcích 1–3 nastavení proměnných, ve kterých se ukládá odkaz na vzdálený Git repozitář, název složky, do níž se obsah má klonovat, a název složky, do které se má kopírovat výstup čili vygenerované HTML. Dále se skript na řádku 5 přepíná do složky, v níž se sám nachází — aby skript fungoval vždy, ať je spuštěný ze kteréhokoliv místa v souborovém systému.

V další části skriptu probíhá na řádku 7 kontrola, zda již existuje složka s naklonovaným Git repozitářem. Pokud složka neexistuje, provede se naklonování vzdáleného repozitáře a tím i vytvoření složky.

Třetí část provádí generování statického obsahu. Nejprve se skript přepne do repozitáře, v němž provede příkaz `git pull`, který do složky stáhne poslední změny ze vzdáleného repozitáře: synchronizuje obsah na poslední verzi. Po synchronizaci repozitáře proběhne

---

<sup>14</sup><https://gitea.com/>

<sup>15</sup>Respektuje základní svobody uživatele (<https://www.gnu.org/philosophy/free-sw.html>)

<sup>16</sup><https://git-scm.com/docs/githooks/>

samotné spuštění generátoru, jenž z obsahu vygeneruje statické HTML, které vloží do složky `./public`. Poté na řádcích 12–14 probíhá kopírování nově vygenerovaného obsahu do složky `/srv/www/ucitelonline` včetně nastavení unixových práv souborů na bezpečné hodnoty, které se liší pro složky a pro soubory.

Fig. 6.19: Skript pro automatizované generování obsahu

```
0 #!/bin/sh -e
1 GREPO="https://git.microlab.space/pedf/ucitelonline"
2 GDIR="ucitelonline"
3 WEBROOT="/srv/www/ucitelonline"
4
5 cd $(dirname $0)
6
7 [ ! -d "$GDIR" ] && git clone --recursive "$GREPO" "$GDIR"
8
9 cd "$GDIR"
10 git pull
11 zola build
12 rsync --recursive --delete --checksum \
13     --group --groupmap=*:www-data --chmod=D750,F640 \
14     public/ "$WEBROOT"
```

Skript spoléhá na to, že systém má již předem správně nakonfigurované uživatele, uživatelské skupiny a web server, a že jsou nainstalované potřebné programy Git, Rsync a generátor Zola. Systémový uživatel, pod kterým je vyvolán Git hook, musí být ve skupině `www-data`, nebo v jiné skupině společně s uživatelem, pod nímž je spuštěn web server. Zároveň musí mít uživatel práva pro zápis do cílové složky, tedy kořenové složky web serveru: `/srv/www/ucitelonline`.

Ve většině případů by bylo vhodné klonovat a generovat obsah v dočasné složce, například v `/tmp`, a po zkopírování souborů do složky web serveru opět zdrojové soubory smazat. To se ovšem v této implementaci nehodí, protože repozitář se zdrojovými soubory může být velký a jeho klonování může potenciálně zabrat zbytečné množství času, na rozdíl od příkazu `git pull`, který pouze stáhne změny. Generátor zároveň při

generování zpracuje pouze nutné změny, zatímco po čistém naklonování musí zpracovat celý obsah znovu, což může také trvat dlouho, obzvláště při zpracování mnoha obrázků. V této implementaci se tedy zachováním naklonovaného repozitáře výrazně zkracuje čas celého skriptu.

# 7. Vyhodnocení modelové implementace

V této části práce je shrnuta a zhodnocena modelová implementace z kapitoly 6, a to jak implementace a využití samotného systému, tak jeho rozšíření implementovaných v sekci 6.6.1. Součástí této kapitoly jsou také návrhy pro další rozšíření systému.

Webové stránky vytvořené v rámci této práce jsou k dispozici na adrese <https://ucitelonline.pedf.cuni.cz/> a jejich zdrojový kód je umístěn v Git repozitáři na adrese <https://git.microlab.space/pedf/ucitelonline/>.

## 7.1 Návrhy pro rozšíření systému

V praxi bylo zjištěno, že uživatelé, kteří neznají verzovací systém Git, mají problémy se jej naučit, obzvláště v prostředí, které vyžaduje rychlé zpracování změn. Systém by bylo dobré rozšířit o jednoduchou webovou administraci, jež umožňuje nezkušeným uživatelům jednoduchou práci s obsahem bez nutnosti hledání souborů ve stromové struktuře a znalosti jazyka Markdown. Částečně je tato funkcionality poskytována systémem Gitea, jenž umožňuje jednodušší úpravy provádět přímo v prohlížeči, ovšem uživatel musí stále znát a pracovat s unikátnostmi jazyka Markdown a generátoru Zola.

Skript 6.19 pro automatické generování obsahu ze sekce 6.6.1 je možné rozšířit tak, aby byl schopen pracovat se vstupem z Git hooku či se standardním vstupem *stdin*, který by umožňoval využití skriptu univerzálně pro různé webové stránky, nikoliv jen specificky pro tuto implementaci. Skript by také bylo možné rozšířit o jednoduché příkazy *echo*, které by oznamovaly stav, v němž se skript nachází. Standardní výstup skriptu vyvolaný přes Git hook je přesměrován uživateli, jenž spustil příkaz `git push` a tím i samotný hook. Skript by poté informoval uživatele o tom, zda právě stahuje změny na server, generuje statický obsah, či kopíruje soubory do kořenové složky web serveru. Z důvodu zachování jednoduchosti skriptu nebyly tyto funkcionality implementovány.

## 7.2 Vyhodnocení implementace vlastních rozšíření

Do systému v modelové implementaci byla přidána vlastní rozšíření, tedy filtry pro vkládání souborů a videí do obsahu stránky, viz 6.4. Tyto filtry splňují původní požadavky, avšak jejich použití v obsahu se vymyká původnímu principu jazyka Markdown: že obsah je čitelný i v čistém textu. Pro vyvolání filtrů je třeba vyplňovat jejich různé atributy, což se může zdát nepřehledné někomu, kdo si fungování filtrů neprostudoval. Zároveň je pak obsah nepřenositelný do jiných systémů, které neumí tyto filtry zpracovat a v nichž by se kód pro vyvolání filtrů v takovém případě mohl interpretovat jako čistý text.

# Závěr

Cílem práce bylo analyzovat a zhodnotit generátory statických webových stránek a jimi využívaných značkovacích jazyků. Součástí bylo také vytvoření modelové implementace, její rozšíření, optimalizace a celkové zhodnocení.

Z práce plyne, že statické webové stránky jsou bezpečnější a rychlejší než dynamické webové aplikace. Nehodí se ovšem pro některá specifická paradigmat. Zároveň je práce s nimi bez pomocných nástrojů složitější hlavně pro uživatele, kteří chtějí do obsahu přispívat a kteří zároveň neumí pracovat s verzovacími systémy. Na druhou stranu je jejich údržba méně složitá v roli serverového administrátora.

K určení nejběžnějších webových paradigmat bylo přistoupeno analýzou nejběžněji se vyskytujících typů webových stránek. Pro každé paradigma byly vybrány statické generátory, které jsou ke generování daného typu stránek nejvhodnější, nebo bylo určeno, že dané paradigma není pro statické generátory vhodné.

Na základě obecných a specificky zadaných požadavků byla sestavena taxonomie požadavků pro modelový web, podle kterého byl modelový systém úspěšně vytvořen. Do systému byla dle požadavků implementována i vlastní rozšíření s cílem nahradit chybějící funkcionality vybraného generátoru. Systém se podařilo ze značné části optimalizovat, přičemž byly vytvořeny návrhy na další možnou optimalizaci.

Vyhodnocení modelové implementace popisuje možná vylepšení systému a jeho nedostatky, z nichž nejzásadnější je složitý systém úprav obsahu a nutná znalost značkovacího jazyka Markdown, specifických rozšíření a verzovacího systému Git.

Práce dokazuje, že při vhodné kombinaci značkovacího jazyka, generátoru, verzovacího systému, skriptů pro automatizaci a pomocných nástrojů je možné vytvořit bezpečný a rychlý systém pro generování statických webových stránek. S rapidním vývojem softwaru v této oblasti může být situace v budoucnu velmi odlišná.

# Seznam použité literatury

- CALOMEL. *Webserver Optimization and Bandwidth Saving Tips* [online]. 2017. [cit. 2020-03-23]. Dostupné z: [https://calomel.org/save\\_web\\_bandwidth.html](https://calomel.org/save_web_bandwidth.html).
- CHACON, S. *Why Git is Better than X* [online]. 2009. [cit. 2020-03-26]. Dostupné z: <http://z.github.io/whygitisbetter/>.
- CHADBURN, M. – LAHAV, G. *How slow websites damage publishers revenue* [online]. 2016. [cit. 2020-02-15]. Dostupné z: <https://web.archive.org/web/20180929125709/http://engineroom.ft.com/2016/04/04/a-faster-ft-com/>.
- CIMPANU, C. *How Static Site Generators Work* [online]. 2015. [cit. 2020-03-16]. Dostupné z: <https://news.softpedia.com/news/How-Static-Site-Generators-Work-482007.shtml>.
- CVE DETAILS. *PHP : Vulnerability Statistics* [online]. 2020a. [cit. 2020-03-31]. Dostupné z: <https://www.cvedetails.com/product/128/PHP-PHP.html>.
- CVE DETAILS. *Wordpress : Vulnerability Statistics* [online]. 2020b. [cit. 2020-03-29]. Dostupné z: <https://www.cvedetails.com/product/4096/Wordpress-Wordpress.html>.
- DOMINIK, C. *Technical description in 24 words* [online]. 2008. [cit. 2020-04-15]. Dostupné z: <https://orgmode.org/worg/org-quotes.html>.
- GOODGER, D. *Problems With StructuredText* [online]. 2012. [cit. 2020-05-01]. Dostupné z: <https://docutils.sourceforge.io/docs/dev/rst/problems.html>.
- GOUY, I. *The Computer Language Benchmarks Game* [online]. 2020. [cit. 2020-03-18]. Dostupné z: <https://benchmarksgame-team.pages.debian.net/benchmarksgame/which-programs-are-fastest.html>.
- GRUBER, J. *Markdown* [online]. 2004. [cit. 2020-02-27]. Dostupné z: <https://daringfireball.net/projects/markdown/>.



- HOFFMAN, B. *Improving Search Rank by Optimizing Your Time to First Byte* [online]. 2013. [cit. 2020-02-12]. Dostupné z: <https://moz.com/blog/improving-search-rank-by-optimizing-your-time-to-first-byte>.
- KHANDELWAL, S. *Critical Bug in WordPress Theme Plugin Opens 200,000 Sites to Hackers* [online]. 2020. [cit. 2020-03-29]. Dostupné z: <https://thehackernews.com/2020/02/themegrill-wordpress-plugin.html>.
- KHANDELWAL, S. *Flaw in Elementor and Beaver Addons Let Anyone Hack WordPress Sites* [online]. 2019. [cit. 2020-03-29]. Dostupné z: <https://thehackernews.com/2019/12/wordpress-elementor-beaver.html>.
- LEONARD, S. Guidance on Markdown: Design Philosophies, Stability Strategies, and Select Registrations. RFC 7764, Internet Engineering Task Force, Mar 2016. Dostupné z: <https://tools.ietf.org/html/rfc7764>.
- MACFARLANE, J. *CommonMark Spec* [online]. 2019. [cit. 2020-03-22]. Dostupné z: <https://spec.commonmark.org/>.
- MARTÍ, V. *A formal spec for GitHub Flavored Markdown* [online]. 2017. [cit. 2020-03-23]. Dostupné z: <https://github.blog/2017-03-14-a-formal-spec-for-github-markdown/>.
- OLŠÁK, P. *Proč nerad používám LaTeX* [online]. 1997. [cit. 2020-03-10]. Dostupné z: <http://petr.olsak.net/ftp/olsak/bulletin/nolatex.pdf>.
- OSSANNA, J. F. – KERNIGHAN, B. W. Troff user's manual. *Bell Laboratories Computing Science Technical Report*. 1977, 54.
- OWASP. OWASP Top Ten 2017. Technical report, OWASP, 2017. Dostupné z: [https://owasp.org/www-pdf-archive/OWASP\\_Top\\_10-2017\\_\(en\).pdf.pdf](https://owasp.org/www-pdf-archive/OWASP_Top_10-2017_(en).pdf.pdf).
- PC MAGAZINE. *Definition of: dynamic Web page* [online]. 2017. [cit. 2020-02-12]. Dostupné z: <https://web.archive.org/web/20170117040526/https://www.pcmag.com/encyclopedia/term/42199/dynamic-web-page>.

PC MAGAZINE. *Definition of: static Web page* [online]. 2020. [cit. 2020-02-12]. Dostupné z: <https://web.archive.org/web/20200223095514/https://www.pcmag.com/encyclopedia/term/static-web-page>.

SCHULTE, E. et al. A Multi-Language Computing Environment for Literate Programming and Reproducible Research. *Journal of Statistical Software*. 1 2012, 46, 3, s. 1–24. ISSN 1548-7660. Dostupné z: <http://www.jstatsoft.org/v46/i03>.

SLANT. *What are the best markup languages?* [online]. 2020. [cit. 2020-02-10]. Dostupné z: <https://web.archive.org/web/20200210061112/https://www.slant.co/topics/589/~best-markup-languages>.

THE ORG MODE DEVELOPERS. *The Org Manual*, 2020.

YAHOO! *Best Practices for Speeding Up Your Web Site* [online]. 2020. [cit. 2020-04-21]. Dostupné z: <https://developer.yahoo.com/performance/rules.html>.

# Seznam ukázek zdrojového kódu

4.1	Příklad podtržení nadpisu ve značkovacím pseudo-jazyce . . . . .	18
6.1	Příklad jednoduché konfigurace v souboru <code>config.toml</code> . . . . .	27
6.2	Základní šablona <code>index.html</code> . . . . .	27
6.3	Základní šablona s filtrem pro přepsání názvu na velká písmena . . . . .	28
6.4	Využití bloků v šabloně z příkladu 6.3 . . . . .	28
6.5	Definice nové šablony <code>section.html</code> rozšiřující šablonu z příkladu 6.4 . . . . .	29
6.6	Vkládání obsahu ze zdrojového Markdown souboru . . . . .	30
6.7	Cyklus pro vypisování všech rodičů v dané větvi navigace . . . . .	31
6.8	Cyklus pro vypisování všech potomků dané stránky do navigace . . . . .	32
6.9	Příklad jednoduchého filtru s jedním atributem . . . . .	33
6.10	Vyvolání vlastního filtru s jedním parametrem . . . . .	33
6.11	Výstup direktivy z příkladu 6.9 . . . . .	34
6.12	Filtr pro vkládání videa s využitím podmínek . . . . .	34
6.13	Vyvolání filtru 6.12 s formátováním na řádky . . . . .	34
6.14	Výstup direktivy z příkladu 6.13 . . . . .	35
6.15	Automatický výpis obrázků s pevně definovanými koncovkami . . . . .	35
6.16	Filtr pro výpis souborů s automatickým hledáním . . . . .	36
6.17	Vyvolání filtru 6.16 . . . . .	37
6.18	Výstup direktivy z příkladu 6.17 . . . . .	37
6.19	Skript pro automatizované generování obsahu . . . . .	43