

MASARYKOVA UNIVERZITA
FAKULTA INFORMATIKY



Portál digitálního kulturního dědictví

BAKALÁRSKA PRÁCA

Peter Koza

Brno, jar 2016

MASARYKOVA UNIVERZITA
FAKULTA INFORMATIKY



Portál digitálního kulturního dědictví

BAKALÁRSKA PRÁCA

Peter Koza

Brno, jar 2016

*Namiesto tejto stránky vložte kópiu oficiálneho podpísaného zadania práce a
prehlásenie autora školského diela.*

Prehlásenie

Prehlasujem, že táto bakalárska práca je mojím pôvodným autorským dielom, ktoré som vypracoval samostatne. Všetky zdroje, pramene a literatúru, ktoré som pri vypracovaní používal alebo z nich čerpal, v práci riadne citujem s uvedením úplného odkazu na príslušný zdroj.

Peter Koza

Vedúci práce: RNDr. Jaroslav Ráček, Ph.D.

Podakovanie

This is the acknowledgement for my thesis,
which can never span multiple paragraphs.

Zhrnutie

Popsat principy digitalizace a publikování oběktů kulturního dědictví. Zaměřit se na tuto problematiku zejména z pohledu regionálních institucí v České republice. Popsat a analyzovat možnosti využití technologie Liferay Portal pro zpracování a publikování záznamů o objektech kulturního dědictví. V praktické části analyzovat, navrhnout, implementovat a nasadit do provozu prototyp konkrétního portálu kulturního dědictví. Analýza a návrh systému bude nad modely UML, implementace bude v Liferay Portál. V rámci práce řešit i vyhledávání nad automaticky rozpoznávanými texty jednotlivých objektů a import dat z prostředí vybrané instituce.

Klíčové slová

keyword1, keyword2, ...

Obsah

1	Portály digitálneho kultúrneho dedičstva	1
1.1	<i>Problémy existujúcich riešení</i>	1
1.1.1	Zobrazenie knižných dát	1
1.1.2	Realizácia vyhľadávania	2
1.2	<i>Poučenie o tvorbe portálov kultúrneho dedičstva</i>	2
2	Analýza	5
2.1	<i>Požiadavky zákazníka</i>	5
2.1.1	Kategórie dokumentov	5
2.1.2	Dokumenty archívnych fondov a múzejných zbierok	5
2.1.3	Dokumenty knižničných fondov podľa štandardov NDK	6
2.1.4	Typy dokumentov	6
2.1.5	Vstup dokumentov	7
2.1.6	Príprava dokumentov k sprístupneniu	7
2.1.7	Operácie nad zverejnenými dokumentmi	8
2.2	<i>Návrh riešenia</i>	9
2.2.1	Grafický vzhľad	9
3	Technológie	11
3.1	<i>Indexácia</i>	11
3.1.1	Výber vyhľadávacieho nástroja	11
3.1.2	Elastic	11
3.1.3	Ďalšie možnosti vyhľadávania	12
4	These are	15
4.1	<i>the available</i>	15
4.1.1	sectioning	15
5	Floats and references	17
6	Mathematical equations	19
7	We have several FONTS at disposal	21
8	Inserting the bibliography	23

9	Inserting the index	25
	Register	27
A	An appendix	29

Zoznam tabuliek

5.1 A weather forecast 18

Zoznam obrázkov

- 5.1 The logo of the Masaryk University at 40 mm 17
- 5.2 The logo of the Masaryk University at $\frac{2}{3}$ and $\frac{1}{3}$ of text
width 18

1 Portály digitálneho kultúrneho dedičstva

Zámerom tejto iniciatívy je vytvoriť webový portál, ktorého prostredníctvom bude zaisťovaná príprava a samotné sprístupnenie digitálneho obsahu vybraných fondov pamäťových inštitúcií pôsobiacich na území Českej republiky širokej verejnosti a odborným bádateľom. Ďalej bude umožňovať on-line úpravu metadát a možnosť priloženia ďalších materiálov ktoré s dokumentom súvisia. Web by mal užívateľa zaujať a vytvoriť dojem jednoducho použiteľnej interaktívnej stránky aj u používateľa, ktorého daná tématika nezaujíma.

1.1 Problémy existujúcich riešení

Digitálny obsah sa primárne skladá z kníh. Problémom existujúcich nástrojov na prehliadanie tohto obsahu je nepríťažlivosť pre užívateľa.

1.1.1 Zobrazenie knižných dát

Tabuľkové zobrazenie knižných záznamov nespĺňa požiadavky modernej webovej aplikácie. Príkladom je katalóg Slezského zemského múzea¹. Vyhľadávanie je realizované jedným formulárom. Výsledky sú zobrazované jednoduchou tabuľkou, ktorá poskytuje minimum informácií o dokumente a jeho uložení. V tabuľke sú ako prvé uvedené informácie Dok a Sign, ktorých názvy sú neintuitívne a teda pre koncového užívateľa pri prvom prístupe na web nepodstatné. Ďalej sú zobrazené údaje autor, názov, časť, rok a počet. Zvyšok údajov je zobrazený až po kliknutí na názov dokumentu. Ostatné atribúty sú neaktívne. V detaile dokumentu chýbajú užívateľsky príťažlivé prvky, ako napríklad mapa uloženia, diskusia k dielu, možnosť rezervácie alebo zobrazenie skenovaných obrázkov. Aj napriek tomu, že tento web podáva hodnoverné informácie o dokumentoch, nie je možné ho označiť za užívateľsky príťažlivý a po prvom použití nenavádza užívateľa k ďalšej návšteve.

1. Vid' <http://knihovna.szmo.cz/katalog/>.

1.1.2 Realizácia vyhľadávania

Na stránke systému Kramerius² je sprístupnené vyhľadávanie dokumentov Národnej knižnice Českej republiky. Užívateľské rozhranie je oproti predošlému webu obohatené o funkciu hľadania v celom texte, ktorá umožňuje jednoduchšie zoznámenie sa s funkciami aplikácie. Pokročilé vyhľadávanie je možné zobrazíť a skryť pridaným tlačidlom. Priamo na úvodnej stránke sú umiestnené príklady skenovaných dokumentov, ktoré odkazom vedú na detail diela. Zobrazenie výsledkov vyhľadávania je riešené jednoduchým zoznamom. Tento prístup prináša vysokú mieru neprehľadnosti, pretože užívateľ nemôže porovnať hodnoty, či zoradiť výsledky na základe jednotlivých atribútov. Detail obsahuje stručné informácie o diele a náhľad strán, ktorý ale nie je dostupný bez inštalácie zásuvného modulu.

1.2 Poučenie o tvorbe portálov kultúrneho dedičstva

Z analýzy existujúcich nástrojov na prehliadanie kultúrneho dedičstva bolo zistené že pri návrhu aplikácie by sme sa mali zamerať na nasledujúce oblasti:

- vyhľadávanie
 - umožniť hľadanie v celom texte
 - pridať prvky pre oživenie ako napr. mapa alebo rezy
- zobrazenie výsledkov
 - vytvoriť prehľadnú tabuľku obsahujúcu relevantné atribúty
 - v prípade použitia mapy poskytnúť možnosť jednoduchého prepnutia
- detail diela
 - pridať užívateľsky príťažlivé prvky
 - * mapa umiestnenia
 - * diskusia k dielu

2. Viď <http://kramerius.nkp.cz/kramerius/Welcome.do>.

1. PORTÁLY DIGITÁLNEHO KULTÚRNEHO DEDIČSTVA

- * zobrazenie náhľadu dokumentov
- * rezervácia knihy

2 Analýza

2.1 Požiadavky zákazníka

Po komunikácii so zákazníkom na AWS boli zistené požiadavky zákazníka zobrazené diagramom na obrázku 546846

Use case diagram jak hovado

2.1.1 Kategórie dokumentov

Dokumenty , ktoré sú predmetom prípravy a zverejnenia prostredníctvom portálu digitálneho kultúrneho dedičstva(ďalej už len PD) sú rozdelené do dvoch základných kategórií:

- dokumenty archívnych fondov a múzejných zbierok
- dokumenty knižničných fondov podľa štandardov NDK¹

Pre vyššie uvedené jednotlivé kategórie je nutné použiť v PD rozdielne pracovné postupy a to hlavne v prípravnej fáze pred zverejnením. Dôvodom je samotná podstata dokumentu z pohľadu toho, či ide o dokumenty, ktoré sú v nemennej podobe (monografie, periodiká) alebo dokumenty, u ktorých sa v čase môžu meniť predovšetkým ich popisné informácie. Pre každú kategóriu sú taktiež definované odlišné dátové štruktúry metadát. Predpokladaný objem dokumentov spravovaných PD je 100 000 jednotlivých dokumentov, 1 500 000 strán prepočítaných na formát A4 a približne 20 terabajtov obrazových dát.

2.1.2 Dokumenty archívnych fondov a múzejných zbierok

Druhy dokumentov:

- fotografie, negatívy
- zväzky - kroniky, zápisy, katalógy
- voľné archy - dokumenty, plagáty, plány
- staré výtlačky, spevníky

1. Národní digitální knihovna

2. ANALÝZA

- kartografický materiál
- filmy, zvukové nahrávky
- zbierkové predmety

Jedná sa o dokumenty, ktoré vznikli digitalizáciou vybraných archívnych fondov a múzejných zbierok. Z hľadiska nemennosti týchto dokumentov sa jedná o dokumenty premenlivej povahy. Popisné metadáta z KDJ² pre PD u tejto kategórie preto nie je možné použiť. Jedná sa teda o dokumenty, u ktorých je nutné ručné doplnenie metadát.

2.1.3 Dokumenty knižničných fondov podľa štandardov NDK

Druhy dokumentov:

- monografie - jednodielne alebo viacdielne knižné dokumenty
- periodiká - pravidelne vychádzajúce výtlačky

Jedná sa o dokumenty prevažne nemennej povahy. Predpokladá sa teda, že ako uložené digitálne obrazy, tak metadáta majú konečný a nemenný stav. Všetky tieto dokumenty sú uložené v úložisku KDJ vo formáte SIP. KDJ bude pre tento typ dokumentov zdrojovým poskytovateľom informácií pre PD.

2.1.4 Typy dokumentov

Z hľadiska typu súborov, v ktorých sú digitálne dokumenty uložené, se jedná o dokumenty:

- obrazové
- videá
- zvukové záznamy
- vo formáte XML(popisné metadáta)
- vo formáte PDF

2. Krajská digitalizační jednotka

2.1.5 Vstup dokumentov

Táto časť PD bude zabezpečovať riadený import dát pre prípravu sprístupnenia dokumentov. Vstupné mechanizmy budú rešpektovať potreby jednotlivých kategórií dokumentu. Po importe budú dáta pripravených dokumentov uložené do zverejňovacej databázy. Princíp predávania dokumentov, detailný popis rozhrania a jeho dátová štruktúra bude súčasťou implementačnej analýzy. Zákazník špecifikoval nasledujúce podmienky:

- PD bude umožňovať import viacerých dokumentov súčasne.
- Pri zahájení prípravy vstupu dokumentu do PD bude prevedená kontrola duplicity s predchádzajúcimi importovanými dokumentmi.
- Behom importu dokumentu do PD systém priradí jeho vlastníka – osobu kompletne zodpovednú za celé sprístupnenie a zverejnenie. Vlastníkom se stáva užívateľ PD, ktorý import úspešne dokončil. Zdrojom dát pre import bude mimo iných webové rozhranie PD a zdieľané diskové zložky unikátne pre každého administrátora.
- PD bude disponovať funkciou pre import a aktualizáciu popisných metadát a prípadných obrazových dát zo systému JANUS2000 (evidenčný systém okresných archívov) a systému VISMO (dáta projektu MG Vysočiny on-line).

2.1.6 Príprava dokumentov k sprístupneniu

Táto časť PD bude zahŕňať proces sprístupnenia dokumentov a možnosť ich úpravy pred zverejnením. Zákazník špecifikoval nasledujúce funkcie:

- PD po prihlásení do časti pre prípravu dokumentu ponúkne oprávnenému užívateľovi zoznam všetkých jeho dokumentov s informáciou o ich stave a možnosťou zverejnenia dokumentu pre všetkých užívateľov.
- PD umožní oprávnenému užívateľovi zobrazíť náhľad upraveného dokumentu, znázorňujúci stav, v ktorom bude dokument zverejnený.

2. ANALÝZA

- PD umožňuje oprávnenému užívateľovi kedykoľvek zrušiť zverejnenie dokumentu.
- Oprávnený užívateľ, vlastník dokumentu, má na karte otvoreného dokumentu k dispozícii históriu úprav.
- Portál PD umožní oprávnenému užívateľovi možnosť úpravy popisných dát podľa úrovne jeho oprávnení.
- Štruktúra dokumentu bude navrhnutá podľa štandardu Dublin Core.
- Všetky dôležité akcie PD ako sprístupnenie, úprava alebo zmazanie dokumentu budú sprevádzané notifikačným e-mailom pre všetkých dotknutých užívateľov.

2.1.7 Operácie nad zverejnenými dokumentmi

Táto časť portálu obsahuje funkcie dostupné pre širokú verejnosť, tzn. registrovaných a neregistrovaných užívateľov. Zákazník špecifikoval nasledujúce funkcie, dostupné všetkým užívateľom:

- PD bude poskytovať možnosť vyhľadávania v celom texte.
- PD bude poskytovať možnosť vyhľadávania podľa konkrétnych atribútov.
- PD umožní export obsahu príslušného dokumentu do PDF súboru.

Nasledujúce funkcie budú dostupné všetkým registrovaným užívateľom:

- Možnosti pre užívateľský popis príslušného dokumentu. Ide o diskusiu k dielu a prepis nerozpoznaného textu dokumentu.
- PD umožní rezerváciu diela pre vypožičanie.

2.2 Návrh riešenia

Na základe analýzy existujúcich nástrojov na prehliadanie digitálnych fondov a komunikácie so zákazníkom bol vytvorený konečný návrh projektu. Pre implementáciu PD bude využité open source portálové riešenie *Liferay* vo verzii 6.2.2 *community edition GA3*. Prípadný prechod na vyššiu verziu je možný prostredníctvom štandardných postupov pre portál *Liferay* a nieje zahrnutý v rozsahu tohto projektu. Zaznamenávanie operácií v PD bude zabezpečené pomocou frameworku pre Java aplikácie *Log4j*. Výstupné hlášky sú štandardne zaznamenávané do súboru na aplikačnom serveri, na ktorom aplikácia beží. Jednotlivé hlášky budú rozdelené do úrovní. Nastavovanie týchto úrovní bude možné prostredníctvom administratívneho rozhrania portálu *Liferay*. Administrácia zverejňovania digitálneho obsahu bude dostupná všetkým oprávneným užívateľom. Vyhľadávanie bude dostupné všetkým užívateľom bez ohľadu na ich identitu.

2.2.1 Grafický vzhľad

Grafický vzhľad aplikácie bude meniteľný pomocou špeciálnych zásuvných modulov podporovaných portálom *Liferay* (tzv. témy). Zmena vzhľadu vyžaduje vývoj alebo úpravu zásuvného modulu.

3 Technológie

3.1 Indexácia

3.1.1 Výber vyhľadávacieho nástroja

Pre potreby indexácie boli v analytickej fáze zvažované tieto možnosti:

- *Elastic* – vyhľadávací nástroj postavený na Java knižnici *Apache Lucene*
- *Sphinx* – vyhľadávací nástroj vytvorený v jazyku C++

Po konzultácii s vývojovým tímom sa nakoniec rozhodlo pre *Elastic* z dôvodu vyššej spoľahlivosti a jednoduchého používania.

3.1.2 Elastic

Elastic je nástroj na indexáciu dát. Jeho primárnym účelom je vyhľadávanie nad relatívne veľkým množstvom textu. Oproti databáze ponúka oveľa rýchlejšiu odpoveď a rozvinutejšie možnosti štruktúrovaného ukladania dát.

Aplikácie využívajúce Elastic s ním komunikujú pomocou HTTP protokolu a dotazy ako aj odpovede sú vo formáte JSON. Príklad dotazu na Elastic:

```
1 GET
2 {
3   "query" : {
4     "term" : { "gender": "female" },
5     "range" : {
6       "age" : {
7         "from" : 20,
8         "to" : 30
9       }
10    }
11  }
```

Tento dotaz vyhľadá všetky záznamy s uvedenými parametrami. *Term* je označenie výrazu, ktorý vyžaduje presný obsah určitého atribútu.

3. TECHNOLOGIE

Range vyhľadá všetky záznamy ktoré sa nachádzajú v určenom rozmedzí. V tomto prípade hľadáme ženu od 20 do 30 rokov. Odpoveďou na takýto dotaz by mohlo byť napríklad:

```
1 "hits":{
2   "total" : 1,
3   "hits" : [
4     { "_index" : "pdkd",
5       "_type" : "person",
6       "_id" : "1",
7       "_source" : {
8         "name" : "Kim",
9         "age" : 22,
10        "gender" : "female"
11      }
12   }
13 ]
14 }
```

Výsledkom dotazu je teda 1 záznam.

3.1.3 Ďalšie možnosti vyhľadávania

Radenie Poradie vrátených záznamov môže byť vytvorené na základe určitého atribútu, relevancie výsledku alebo kombinácie týchto faktorov.

Score Ďalšou zaujímavou funkciou elasticu je *relevance scoring*. Po tom čo získame zoznam vyhovujúcich záznamov, je potrebné ich zoradiť podľa relevancie. V prípade kombinovaného dotazu s logickými spojkami neobsahuje každý záznam všetky podmienky. Skóre relevancie sa počíta na základe troch faktorov:

1. **Frekvencia výskytu hľadaného výrazu v zázname** - čím viac je v zázname zmienená, tým je záznam podstatnejší.
2. **Inverzná frekvencia hľadaného výrazu** – relevancia výrazu sa znižuje v závislosti od množstva záznamov obsahujúcich tento výraz.
3. **Rozsah atribútu , v ktorom bol výraz nájdený** – čím je obsah atribútu dlhší, tým je relevancia tohto výskytu nižšia

Aggregations Výsledky vyhľadávania je niekedy potrebné zlučovať do skupín na základe určitého parametru. *Elastic* ponúka funkciu *Aggregations*. Pokiaľ je teda v dotaze špecifikovaný atribút, na základe ktorého sa majú záznamy zoskupovať je výsledkom objekt obsahujúci niekoľko skupín ďalších objektov, ktoré môžu predstavovať samotné záznamy alebo ďalšie skupiny zjednotení. Vďaka formátu JSON je možné tieto objekty serializovať. V prípade objektových jazykov ako Java je potom jednoduché poskladať štruktúry, s ktorými jazyk pracuje pomocou knižnice na deserializáciu.

4 These are

4.1 the available

4.1.1 sectioning

commands.

Paragraphs and

subparagraphs are available as well. Inside the text, you can also use unnumbered lists,

- such as
- this one
 - and they can be nested as well.
 - » You can even turn the bullets into something fancier,
 - § if you so desire.

Numbered lists are

1. very
 - (a) similar

and so are description lists:

Description list A list of terms with a description of each term

The spacing of these lists is geared towards paragraphs of text. For lists of words and phrases, the paralist package offers commands

- that
 - are
 - * better
 - suited
- 1. to
 - (a) this
 - i. kind of
 - A. content.

4. THESE ARE

The `amsthm` package provides the commands necessary for the typesetting of mathematical definitions, theorems, lemmas and proofs.

Theorem 4.1.1. *This is a theorem that offers a profound insight into the mathematical sectioning commands.*

Theorem 4.1.2 (Another theorem). *This is another theorem. Unlike the first one, this theorem has been endowed with a name.*

Lemma 4.1.3. *Let us suppose that $x^2 + y^2 = z^2$. Then*

$$\left\langle u \left| \sum_{i=1}^n F(e_i, v) e_i \right. \right\rangle = F\left(\sum_{i=1}^n \langle e_i | u \rangle e_i, v\right). \quad (4.1)$$

Dôkaz. $\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$ □

Corrolary 4.1.4. *This is a corrolary.*

Remark. This is a remark.

5 Floats and references

The logo of the Masaryk University is shown in Figure 5.1 and Figure 5.2 at pages 17 and 18. The weather forecast is shown in Table 5.1 at page 18. The following chapter is Chapter 6 and starts at page 19. Items 3, 3b, and 3(c)iv are starred in the following list:

1. some text
2. some other text
3. ★
 - (a) some text
 - (b) ★
 - (c) some other text
 - i. some text
 - ii. some other text
 - iii. yet another piece of text
 - iv. ★
 - (d) yet another piece of text
4. yet another piece of text

If your reference points to a place that has not yet been typeset, the `\ref` command will expand to `??` during the first run of `pdflatex` `fi-pdflatex.tex` and a second run is going to be needed for the references to resolve. With online services – such as Overleaf – this is performed automatically.



Obr. 5.1: The logo of the Masaryk University at 40 mm



Obr. 5.2: The logo of the Masaryk University at $\frac{2}{3}$ and $\frac{1}{3}$ of text width

Day	Min Temp	Max Temp	Summary
Monday	13°C	21°C	A clear day with low wind and no adverse current advisories.
Tuesday	11°C	17°C	A trough of low pressure will come from the northwest.
Wednesday	10°C	21°C	Rain will spread to all parts during the morning.

Tabulka 5.1: A weather forecast

6 Mathematical equations

T_EX comes pre-packed with the ability to typeset inline equations, such as $e^{ix} = \cos x + i \sin x$, and display equations, such as

$$\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

L^AT_EX defines the automatically numbered equation environment:

$$\gamma Px = PAx = PAP^{-1}Px. \quad (6.1)$$

The package amsmath provides several additional environments that can be used to typeset complex equations:

1. An equation can be spread over multiple lines using the `multline` environment:

$$\begin{aligned} a + b + c + d + e + f + b + c + d + e + f + b + c + d + e + f \\ + f + g + h + i + j + k + l + m + n + o + p + q \end{aligned} \quad (6.2)$$

2. Several aligned equations can be typeset using the `align` environment:

$$a + b = c + d \quad (6.3)$$

$$u = v + w + x \quad (6.4)$$

$$i + j + k + l = m \quad (6.5)$$

3. The `alignat` environment is similar to `align`, but it doesn't insert horizontal spaces between the individual columns:

$$a + b + c + d \quad = 0 \quad (6.6)$$

$$e + f + g = 5 \quad (6.7)$$

4. Much like chapter, sections, tables, figures, or list items, equations – such as (6.8) and (My equation) – can also be labeled and referenced:

$$b_{11}x_1 + b_{12}x_2 + b_{13}x_3 \quad = y_1, \quad (6.8)$$

$$b_{21}x_1 + b_{22}x_2 \quad + b_{24}x_4 = y_2. \quad (\text{My equation})$$

5. The `gather` environment makes it possible to typeset several equations without any alignment:

$$\psi = \psi\psi, \tag{6.9}$$

$$\eta = \eta\eta\eta\eta\eta\eta, \tag{6.10}$$

$$\theta = \theta. \tag{6.11}$$

6. Several cases can be typeset using the `cases` environment:

$$|y| = \begin{cases} y & \text{if } z \geq 0, \\ -y & \text{otherwise.} \end{cases} \tag{6.12}$$

For the complete list of environments and commands, consult the `amsmath` package manual¹.

1. See <http://mirrors.ctan.org/macros/latex/required/amslatex/math/amslldoc.pdf>. The `\url` command is provided by the package `url`.

7 We have several FONTS *at disposal*

The serifed roman font is used for the main body of the text. *Italics are typically used to denote emphasis or quotations.* The teletype font is typically used for source code listings. The **bold**, SMALL-CAPS and sans-serif variants of the base roman font can be used to denote specific types of information.

We can also change the font size, although it is usually not necessary.

A wide variety of mathematical fonts is also available, such as:

$ABC, \mathcal{ABC}, \mathbf{ABC}, \text{ABC}, \textit{ABC}, \text{ABC}$

By loading the amsfonts packages, several additional fonts will become available:

$\mathfrak{ABC}, \mathbb{ABC}$

Many other mathematical fonts are available¹.

1. See <http://tex.stackexchange.com/a/58124/70941>.

8 Inserting the bibliography

After loading the `biblatex` package and linking a bibliography database file to the document using the `\addbibresource` command, you can start citing the entries. This is just dummy text [**inbook-full**] lightly sprinkled with citations [**incollection-full**]. Several sources can be cited at once [**whole-collection**, **manual-minimal**, **manual-full**]. **inbook-full** was written by **inbook-full** in **inbook-full**. We can also produce **inbook-full** or (**inbook-full**, **inbook-full**). The full bibliographic citation is: **inbook-full**. We can easily insert a bibliographic citation into the footnote¹.

The `\nocite` command will not generate any output, but it will insert its argument into the bibliography. The `\nocite{*}` command will insert all the records in the bibliography database file into the bibliography. Try uncommenting the command and watch the bibliography section come apart at the seams.

When typesetting the document for the first time, citing a work will expand to [**work**] and the `\printbibliography` command will produce no output. It is now necessary to generate the bibliography by running `biber fi-pdflatex.bcf` from the command line and then by typesetting the document again twice. During the first run, the bibliography section and the citations will be typeset, and in the second run, the bibliography section will appear in the table of contents.

The `biber` command needs to be executed from within the directory, where the L^AT_EX source file is located. In Windows, the command line can be opened in a directory by holding down the **Shift** key and by clicking the right mouse button while hovering the cursor over a directory. Select the **Open Command Window Here** option in the context menu that opens shortly afterwards.

With online services – such as Overleaf – all commands are executed automatically.

1. **inbook-full**.

9 Inserting the index

After using the `\makeindex` macro and loading the `makeidx` package that provides additional indexing commands, index entries can be created by issuing the `\index` command. It is possible to create ranged index entries, which will encompass a span of text. To insert complex typographic material – such as α or \TeX – into the index, you need to specify a text string, which will determine how the entry will be sorted. It is also possible to create hierarchical entries.

After typesetting the document, it is necessary to generate the index by running

```
texindy -I latex -C utf8 -L <locale> fi-pdflatex.idx
```

from the command line, where *<locale>* corresponds to the main locale of your thesis – such as `english`, and then typesetting the document again.

The `texindy` command needs to be executed from within the directory, where the \LaTeX source file is located. In Windows, the command line can be opened in a directory by holding down the `[Shift]` key and by clicking the right mouse button while hovering the cursor over a directory. Select the `[Open Command Window Here]` option in the context menu that opens shortly afterwards.

With online services – such as Overleaf – the commands are executed automatically, although the locale may be erroneously detected, or the `makeindex` tool (which is only able to sort entries that contain digits and letters of the English alphabet) may be used instead of `texindy`. In either case, the index will be ill-sorted.

Register

α , 19

dummy text, 19

T_EX, 19

vehicles

 speed cars, 19

 trucks, 19

A An appendix

Here you can insert the appendices of your thesis.