

Marcus Cramer, Markus Heuchert

An Investigation in the Springleaf Kaggle Challenge

Seminar Thesis

in the context of the seminar “Applied Machine Learning”

at the Chair for Information Systems and Statistics
(Westphalian Wilhelms-University, Münster)

Principal Supervisor: Prof. Dr. Heike Trautmann

Associate Supervisor:

Tutor: Pascal Kerschke, M.Sc.

Presented by: Marcus Cramer, Markus Heuchert [123456]
Schlossplatz 2
48149 Münster
+49 251 8338100
{m_heuc03, m_cram02}@uni-muenster.de

Submission: 10th January 2017

Any sufficiently advanced technology is indistinguishable from magic.

ARTHUR C. CLARKE

Contents

Figures	V
Tables	VI
Listings	VII
Abbreviations	VIII
Symbols	IX
1 Introduction	1
2 TEMPLATE INTRODUCTION	2
3 TEMPLATE ELEMENTS	3
3.1 Structure	3
3.1.1 Lists	3
3.1.2 Footnotes	3
3.1.3 ToDo-Notes	3
3.2 Formatting Text	3
3.2.1 Colors	4
3.3 Figures	4
3.3.1 Subfigures	5
3.4 Listings	5
3.5 Algorithms	6
3.6 Acronyms and Abbreviations	6
3.6.1 Common Abbreviations	6
3.6.2 Custom Abbreviations	9
3.6.3 Symbols	9
3.7 Citations and Bibliography	9
3.7.1 Misc	10
4 TEMPLATE COMPILING	11
Appendix	12
References	12

Figures

1	Relationship of students and theses	4
2	Exemplary use of subfigures	5

Tables

1	Colors defined by the template	4
2	Common abbreviation macros for English theses	7
3	Common abbreviation macros for German theses	8
4	Commands for printing abbreviations	9

Listings

Listing 1	Euclid's GCD algorithm implemented in Java	5
Listing 2	Commands to compile this document	11
Listing 3	Installing and running Grunt	11

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

1 Introduction

2 TEMPLATE INTRODUCTION

This L^AT_EX template has been developed as an alternative to the well-known Microsoft Word “Becker-Vorlage”. `00_thesis.tex` is the master file.

It is build by Jan Betzing and Dominik Lekse and draws from the DBIS template by Till Haselmann and Florian Stahl, as well as from the IS template by Stephan Dlugosz.

This document is work-in-progress and provides instructions on how to use the template. It does not give advices on scientific writing.

Please feel free to contribute to this template. Members of the WWU Münster can request access to the template by contacting the author at jan.betzing@ercis.uni-muenster.de. Afterwards you will be able to clone the template from <https://wiwi-gitlab.uni-muenster.de/lisis/isthesis.git>, and create push-requests with their new features.

TODO

- Configuration switch for having `\chapter{ }` begin on a new page
- Replace `kvoptions` with `pgfkeys`

3 TEMPLATE ELEMENTS

This chapter gives examples on what you can do with this template. It’s just a brief overview. Please consult the common sources on how to write scientific documents and documents with L^AT_EX.

3.1 Structure

This template provides three structural levels that appear in the table of contents, viz., `\chapter`, `\section`, and `\subsection`. Chapters will always start on a new page. Additionally, you can use `\subsubsection` and `\paragraph` as non-hierarchical means to structure your thesis.

3.1.1 Lists

You can use the default L^AT_EX functions for writing lists, viz., `\enumerate` for numbered lists and `\itemize` for bullet point lists. Again, the `\subsubsection` and `\paragraph` can be used as structural elements, e.g., when listing definitions of terms.

3.1.2 Footnotes

Footnotes are contiguously numbered throughout the whole document. Use the `\footnote{text}` command. They appear on the page their reference is on ¹. Footnotes have to be placed without whitespace behind the word and within the sentence boundaries, i.e., before the period.

3.1.3 ToDo-Notes

You can use ToDo notes using the `\todo{text}` command. Please make sure to remove any ToDo notes before handing in your thesis!

ToDo: Remove me before publishing

3.2 Formatting Text

L^AT_EX provides `\textit{text}` for *italics*, `\textbf{text}` for **bold face**, `\texttt{text}` for typewriter, `\textsc{text}` for SMALL CAPS, `\underline{text}` for underline. Additionally, the template provides `\texthl{text}` for highlighted text. Please remove any highlighted text before handing in your thesis!

Please use the `\enquote{text}` command for “direct quotes”.

¹ This is an exemplary footnote.

3.2.1 Colors

This template comes with the colors defined in the Corporate Designs (CDs) of the ERCIS and WWU. Tab. 1 lists the color names. You can apply them to text by using the `\textcolor{color name}{text}` command.

Color Name	Result
ercis-black	Exemplary Text and 0123456789
ercis-grey	Exemplary Text and 0123456789
ercis-red	Exemplary Text and 0123456789
ercis-lightred	Exemplary Text and 0123456789
ercis-blue	Exemplary Text and 0123456789
ercis-darkblue	Exemplary Text and 0123456789
ercis-cyan	Exemplary Text and 0123456789
ercis-orange	Exemplary Text and 0123456789
ercis-green	Exemplary Text and 0123456789
wwu-black	Exemplary Text and 0123456789
wwu-green	Exemplary Text and 0123456789
wwu-lightgreen	Exemplary Text and 0123456789
wwu-blue	Exemplary Text and 0123456789
wwu-lightblue	Exemplary Text and 0123456789

Table 1 Colors defined by the template

3.3 Figures

The `figure` environment is wrapped around images. These images should either be included as PDF-file via `\includegraphics`, or created via *TikZ/PGF*. For included images, make sure to use high-resolution images, preferably vector images.

Figures float, i.e., they do not necessarily appear at exact the same position you have defined them. Make sure to set a *caption* and an optional *label* as figure parameters.



Figure 1 Relationship of students and theses

3.3.1 Subfigures

Sometimes it might be handy to contrast figures, i.e., by placing them next to each other. The template uses the *subcaption* package to provide subfigures. The following example contains two figures, where each subfigure has its own `\label` and `\caption`. Additionally, the whole figure has its own *caption* and *label*. That means, you can reference subfigures fig. 2a and fig. 2. Only the whole figure will be listed in the table of figures.

Subfigures are not limited to images, but may also include listings or tables. Fig. 2 shows a sample database query expressed in Structured Query Language (SQL) (fig. 2a) and as query plan in relational algebra (fig. 2b).

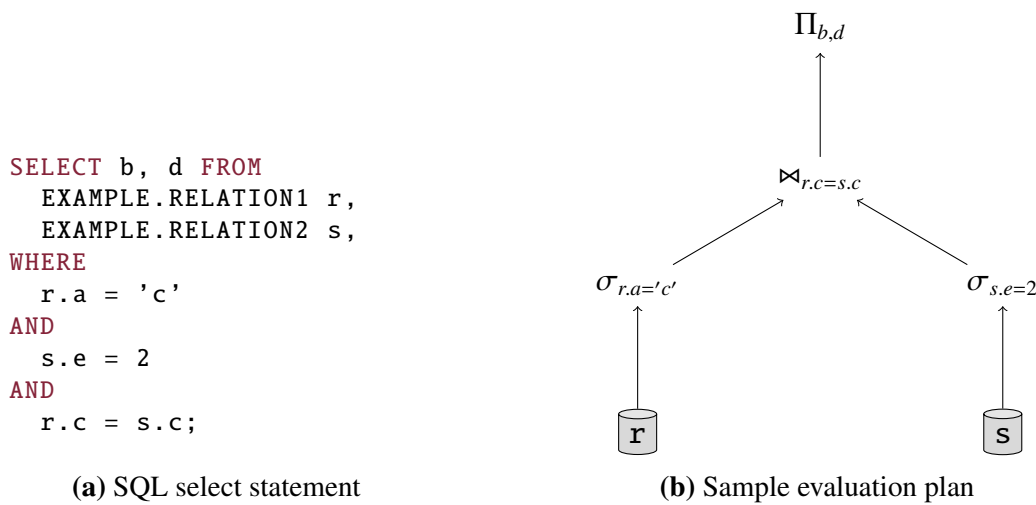


Figure 2 Exemplary use of subfigures

3.4 Listings

You can use listings to typeset source code. This template uses the *listings* package. Wrap code inside the `lstlisting` environment and set the *language* (e.g., Java, SQL), *caption*, and optional *label* parameters. If the source code highlighting highlights the wrong keywords or misses keywords, use the *deletekeywords* resp. *morekeywords* parameters. Consult the package documentation for further information.

```

public class Euclid {

    public static int gcd(int p, int q) {
        if (q == 0) return p;
        else return gcd(q, p % q);
    }

}
        
```

Listing 1 Euclid's GCD algorithm implemented in Java

3.5 Algorithms

Some users might require specifying algorithms. This template uses the *algorithm*, *algorithmicx*, and *algpseudocode* packages. Consult the respective manuals for further information. Algorithms do not appear in a table at the beginning of the document, i.e., there is no list of algorithms.

Algorithm 1 Euclid’s GCD algorithm in pseudocode

Require: nonnegative integer a , nonnegative integer b

function EUCLID(a, b)

if $b = 0$ **then**

 ▷ comment

 return a ;

else

 return EUCLID($b, a \bmod b$);

3.6 Acronyms and Abbreviations

This template provides comprehensive support for acronyms and abbreviations. The template uses the *glossaries* package. Please do only define abbreviations and symbols that are uncommon. That means, common abbreviations such as “e.g.” or “i.e.” should not be listed. Abbreviations and symbols are sorted automatically by their label.

3.6.1 Common Abbreviations

Please note that each full stop in a common abbreviation should be followed by a non-breaking space. This template comes with a variety of macros for common abbreviations, that can be used throughout your theses. The macros differ for English and German theses. Please see the tables below.

Command	Result
<code>\apprx</code>	approx.
<code>\as</code>	a.s.
<code>\cf</code>	cf.
<code>\eg</code>	e.g.
<code>\Eg</code>	E.g.
<code>\esp</code>	esp.
<code>\etal</code>	et al.
<code>\fig</code>	fig.
<code>\Fig</code>	Fig.
<code>\ie</code>	i.e.
<code>\Ie</code>	I.e.
<code>\iid</code>	i.i.d.
<code>\p{4711}</code>	p. 4711
<code>\pf{4711}</code>	p. 4711 f.
<code>\pp{11--47}</code>	pp. 11–47
<code>\resp</code>	resp.
<code>\sect</code>	sec.
<code>\tab</code>	tab.
<code>\Tab</code>	Tab.
<code>\viz</code>	viz.
<code>\wrt</code>	w.r.t.

Table 2 Common abbreviation macros for English theses

Command	Result	Command	Result
<code>\aaO</code>	a. a O.	<code>\oE</code>	o. E.
<code>\Abb</code>	Abb.	<code>\oEdA</code>	o. E. d. A.
<code>\bspw</code>	bspw.	<code>\OEdA</code>	O. E. d. A.
<code>\bzgl</code>	bzgl.	<code>\oV</code>	o. V.
<code>\bzw</code>	bzw.	<code>\OV</code>	O. V.
<code>\ca</code>	ca.	<code>\resp</code>	resp.
<code>\dgl</code>	dgl.	<code>\S{123}</code>	S. 123
<code>\dsgl</code>	dsgl.	<code>\Sf{123}</code>	S. 123 f.
<code>\dh</code>	d. h.	<code>\Sff{123}</code>	S. 123 ff.
<code>\etc</code>	etc.	<code>\siehe</code>	s. o.
<code>\eV</code>	e. V.	<code>\sog</code>	sog.
<code>\evtl</code>	evtl.	<code>\sS{123}</code>	s. S. 123
<code>\fs</code>	f. s.	<code>\sSf{123}</code>	s. S. 123 f.
<code>\gdw</code>	g. d. w.	<code>\sSff{123}</code>	s. S. 123 ff.
<code>\ggf</code>	ggf.	<code>\stu</code>	st. u.
<code>\hc</code>	h. c.	<code>\su</code>	s. u.
<code>\iAllg</code>	i. Allg.	<code>\Tab</code>	Tab.
<code>\iBa</code>	i. B. a.	<code>\tw</code>	t. w.
<code>\idR</code>	i. d. R.	<code>\ua</code>	u. a.
<code>\ieS</code>	i. e. S.	<code>\etal</code>	et al.
<code>\inkl</code>	inkl.	<code>\uae</code>	u. ä.
<code>\insb</code>	insbes.	<code>\uAe</code>	u. Ä.
<code>\Prof</code>	Prof.	<code>\uiv</code>	u. i. v.
<code>\Dr</code>	Dr.	<code>\usw</code>	usw.
<code>\PD</code>	PD.	<code>\uU</code>	u. U.
<code>\Ing</code>	Ing.	<code>\va</code>	v. a.
<code>\iV</code>	i. V.	<code>\vgl</code>	vgl.
<code>\iW</code>	i. W.	<code>\Vgl</code>	Vgl.
<code>\iwS</code>	i. w. S.	<code>\vs</code>	v. s.
<code>\Nr{123}</code>	Nr. 123	<code>\zB</code>	z. B.
<code>\nW</code>	n. W.	<code>\zT</code>	z. T.
<code>\oa</code>	o. a.	<code>\zz</code>	zz.
<code>\oAe</code>	o. Ä.	<code>\zzgl</code>	zzgl.
<code>\oae</code>	o. ä.		

Table 3 Common abbreviation macros for German theses

3.6.2 Custom Abbreviations

Custom abbreviations are defined in the `acronyms.tex` file, using the `\newacronym[longplural={<long plural>}, shortplural={<short plural>}]{<label>}{<short>}{<long>}` command. The *longplural* and *shortplural* parameters are optional. The abbreviations are sorted by their labels. The label is furthermore used to reference the abbreviations in your text. You can do so using commands listed in tab. 4. In most cases, you just use `\gls{<label>}`. On the first occurrence, the full version is displayed, e.g., European Research Center for Information Systems (ERCIS). Afterwards, the short version will be displayed, e.g., ERCIS.

You pluralize your abbreviation by adding a `pl` to the resp. command. This will add a small `s` to the abbreviation, e.g., CDs. Tab. 4 shows custom short and long plural versions of the abbreviation KMU. You might need this esp. for more complex German abbreviations that do not have a “s” plural form.

Command	Result
<code>\gls{<label>}</code>	<code>\acrfull</code> on first occurence, <code>\acrshort</code> otherwise
<code>\glspl{<label>}</code>	<code>\acrfullpl</code> on first occurence, <code>\acrshortpl</code> otherwise
<code>\acrshort{<label>}</code>	KMU
<code>\acrshortpl{<label>}</code>	KMUen
<code>\aclong{<label>}</code>	Kleines und Mittleres Unternehmen
<code>\aclongpl{<label>}</code>	Kleine und Mittlere Unternehmen
<code>\acrfull{<label>}</code>	Kleines und Mittleres Unternehmen (KMU)
<code>\acrfullpl{<label>}</code>	Kleine und Mittlere Unternehmen (KMUen)

Table 4 Commands for printing abbreviations

Only referenced abbreviations will be added to the list of abbreviations.

3.6.3 Symbols

If required, you can define symbols in the `symbols.tex` file, using the `\addsymboltolist{<symbol>}{<label>}{<name>}` command. The symbols are sorted by their labels. Please note, regardless of using the symbols in the text, all symbols defined in the symbols file will be output to the list of symbols.

3.7 Citations and Bibliography

This template uses BibTeX for bibliographies. It comes with the MISQ style that takes care of proper formatting and sorting of your references. Of course, you have to maintain a clean `.bib`

file that caters all necessary attributes. References will appear in the alphabetical order of the surname of the first author. In case of several works by the same author, they are sorted by year.

Citing in the text is done with the `\citep[<before>][<after>]{<citekey>}` command. Citations without parenthesis are done with `\cite{<citekey>}`. You can reference authors with `\citeauthor{<citekey>}`. However, we suggest typesetting authors in SMALL CAPS, e.g., ? is one father of Business Process Management (BPM).

Exemplary citations

- BPM is an integral management paradigm for building and running effective and efficient organizations (??).
- A holistic approach to BPM goes beyond process modeling and workflow management systems (?, p. 530).
- See ? for a comprehensive review on BPM best practices.
- ? lists organizational capabilities for BPM (cf. ?, p. 9 f.), while vom BROCKE et al. give principles of good BPM (cf. ?, pp. 530–546).
- Two authors are automatically divided by an “and” in English or an “und” in German, e.g., (?).
- “BPM can provide a solid set of capabilities essential to master contemporary and future challenges” (?, p. 534).

3.7.1 Misc

The name and matriculation number of the student will automatically be displayed on the header of every page when the thesis type *seminar* is selected.

4 TEMPLATE COMPILING

In order to generate a PDF-file from your \TeX -file you have to run the following commands. We assume you have a master file `00_thesis.tex` that you want to typeset.

```
pdflatex 00_thesis
pdflatex 00_thesis
makeglossaries 00_thesis
bibtex 00_thesis
pdflatex 00_thesis
pdflatex 00_thesis
```

Listing 2 Commands to compile this document

Alternatively, you can use your favorite task runner. This thesis comes with a *Grunt* file to kick-start your \LaTeX writing.

When running, Grunt will monitor your thesis and on file changes, the PDF-file is automatically rebuild using the commands from listing 2.

Please make sure to have node.js and the Node Package Manager (NPM) installed. Now you can open a command prompt at the document root and run the commands in listing 3.

```
# Install Grunt via npm (use sudo on Unix-based OS)
npm install -g grunt-cli

# Install Grunt plugins / dependencies
npm install

# Run the Grunt listener
grunt
```

Listing 3 Installing and running Grunt

Appendix

A Some Appendix Section

Appendices provide only two structural levels, viz., `\section`, and `\subsection`.

The numbering of figures, listings, tables, and footnotes is not reset. Thus, it continues as usual in the appendix.

A.1 Some Appendix Subsection

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetur odio sem sed wisi.

Declaration of Authorship

I hereby declare that, to the best of my knowledge and belief, this Seminar Thesis titled “An Investigation in the Springleaf Kaggle Challenge” is my own work. I confirm that each significant contribution to and quotation in this thesis that originates from the work or works of others is indicated by proper use of citation and references.

Münster, 10th January 2017

Marcus Cramer, Markus Heuchert