# **Report Template**

firstname lastname mail@example.org

firstname lastname mail@example.org

#### **ABSTRACT**

A generally good scheme for an abstract is the following:

- The problem?
- Our solution.
- Our solution in detail.
- So what?

(1-2 sentences each). You don't need to provide an abstract during the first 3 phases, as you most likely will not have an answer to all items. It should be included in your phase 4 submission and the final document.

# INTRODUCTION

This template should provide you some hints on preparing the reports. It is based on the template as provided by ACM<sup>1</sup>.

When you submit the report, indicate who is responsible for a particular phase (either add the name to the section heading of the phase or add a section indicating all the phase responsibles).

Remove the irrelevant sections before submission. Your submission for phase 1 should only contain the first section (Problem Statement) and References (and potentially the abstract). Extend the document for the further phases, i.e. the submission for phase 2 should contain section 1 (Problem Statement) and 2 (Data Acquisition) and so on.

Submit the pdf version of your report to all tutors via e-mail:

- sahib.julka@uni-passau.de
- julian.stier@uni-passau.de
- joerg.schloetterer@uni-passau.de

### 1 PROBLEM STATEMENT

Describe your problem statement clearly and short. Your algorithm will most likely not bring world-peace, but solve a particular problem. Narrow down your problem to be as specific as possible. The more specific your problem, the better the chance to find a suitable solution. After a solution is found, the complexity can still be increased. For example "the beer price is too high, we aim to reduce it" is not a very precise problem statement. What is "too high"? Which kind of beer? At which area/country? At which occasion (restaurant, supermarket, special event, etc.)? …?

Besides the problem statement, you should also indicate your approach on how you plan to solve it and how you plan to evaluate your solution in this section.

Most likely, the problem you are going to address is not completely new, but at least a similar problem has previously been

<sup>1</sup>https://www.acm.org/publications/proceedings-template

Data Science Lab, 2019, Uni Passau

firstname lastname mail@example.org

firstname lastname mail@example.org

addressed by others already. Therefore, research how others solved this or similar problems, i.e. get familiar with the state of the art and related work. You may then apply (and adapt/extend) an existing solution to a similar problem to your problem at hand.

# 2 DATA ACQUISITION & PRE-PROCESSING

Make sure to precisely describe what you did and why. The reader of your report should be able to reproduce the steps. Therefore, it is for example not sufficient if you write "we tokenized the sentences". You need to describe how the tokenization was done exactly, i.e. which regular expression or library/method you used with which parameters.

Frameworks/libraries are mentioned in footnotes instead of in the references section. For example "we used the *TweetTokenizer* from the NLTK<sup>2</sup> toolkit with the default parameters to tokenize our tweets".

Reasoning is highly important. It should be obvious to the reader, why you do something in a particular way. The following subsections provide hints on what to include in your report.

(!) required (\*) if it applies to your project

# 2.1 Data acquisition

- source of data (!)
- means of acquisition (!)
- reasoning (!) (i.e. why you chose that kind of data, that source, why you crawled it in that way, etc.)

As the datasets are mostly pre-defined (i.e. in most of the topics you do not need to acquire the data yourself), this part can be short.

# 2.2 Data preprocessing

- filtering / grouping / labeling (\*)
- lemmatization/stemming (\*)
- other (\*)
- statistics on the data (!) (e.g. volume, classes, distributions, correlations, etc.)
- reasoning (!)

# 2.3 Feature engineering

- input x and output y of the system (\*)
- feature extraction (\*)
- feature transformation (\*)
- reasoning (!)

 $<sup>^2</sup> https://www.nltk.org/api/nltk.tokenize.html\\$ 

#### 3 MODEL IMPLEMENATION

# 3.1 Methodology / Proposed solution / Technique

An in-depth description of your solution to the problem. (e.g. input - output, system description, ML techniques, etc.)

# 3.2 Experimental Setup

An in-depth description of the experimental design which enables an objective quantification of the quality of your solution. (e.g. dataset, baselines, metrics, etc.). Might be moved to the next section (Phase 4).

### 4 EVALUATION

### 4.1 Results

The evaluation of your solution by means of the experiments. Figures and statistics provide hard facts about the quality of your solution from different viewpoints.

## 4.2 Discussion

So what? How well could we solve the problem? What are the limitations? Open ends.

# TEMPLATE INFORMATION (REMOVE THIS SECTION IN YOUR REPORT)

Instructions in the following sections are included from the original ACM template sample file. The article template's documentation, available at https://www.acm.org/publications/proceedings-template, has a complete explanation and tips for effective use.

# **Sectioning Commands**

Your work should use standard LATEX sectioning commands: section, subsection, subsection, and paragraph.

### **Tables**

The "acmart" document class includes the "booktabs" package — https://ctan.org/pkg/booktabs — for preparing high-quality tables.

Table captions are placed *above* the table.

Because tables cannot be split across pages, the best placement for them is typically the top of the page nearest their initial cite. To ensure this proper "floating" placement of tables, use the environment **table** to enclose the table's contents and the table caption. The contents of the table itself must go in the **tabular** environment, to be aligned properly in rows and columns, with the desired horizontal and vertical rules. Again, detailed instructions on **tabular** material are found in the ETEX User's Guide.

Immediately following this sentence is the point at which Table 1 is included in the input file; compare the placement of the table here with the table in the printed output of this document.

To set a wider table, which takes up the whole width of the page's live area, use the environment **table**\* to enclose the table's contents and the table caption. As with a single-column table, this wide table will "float" to a location deemed more desirable. Immediately following this sentence is the point at which Table 2 is included in

**Table 1: Frequency of Special Characters** 

Non-English or Math	Frequency	Comments
Ø	1 in 1,000	For Swedish names
$\pi$	1 in 5	Common in math
\$	4 in 5	Used in business
$\Psi_1^2$	1 in 40,000	Unexplained usage

the input file; again, it is instructive to compare the placement of the table here with the table in the printed output of this document.

# **Math Equations**

You may want to display math equations in three distinct styles: inline, numbered or non-numbered display. Each of the three are discussed in the next sections.

Inline (In-text) Equations. A formula that appears in the running text is called an inline or in-text formula. It is produced by the **math** environment, which can be invoked with the usual \begin . . . \end construction or with the short form \$ . . . \$. You can use any of the symbols and structures, from  $\alpha$  to  $\omega$ , available in MTEX [20]; this section will simply show a few examples of in-text equations in context. Notice how this equation:  $\lim_{n\to\infty} x=0$ , set here in in-line math style, looks slightly different when set in display style. (See next section).

*Display Equations.* A numbered display equation—one set off by vertical space from the text and centered horizontally—is produced by the **equation** environment. An unnumbered display equation is produced by the **displaymath** environment.

Again, in either environment, you can use any of the symbols and structures available in LATEX; this section will just give a couple of examples of display equations in context. First, consider the equation, shown as an inline equation above:

$$\lim_{n \to \infty} x = 0 \tag{1}$$

Notice how it is formatted somewhat differently in the **displaymath** environment. Now, we'll enter an unnumbered equation:

$$\sum_{i=0}^{\infty} x + 1$$

and follow it with another numbered equation:

$$\sum_{i=0}^{\infty} x_i = \int_0^{\pi+2} f$$
 (2)

just to demonstrate LATEX's able handling of numbering.

# **Figures**

The "figure" environment should be used for figures. One or more images can be placed within a figure. If your figure contains third-party material, you must clearly identify it as such, as shown in the example below.

Your figures should contain a caption which describes the figure to the reader. Figure captions are placed *below* the figure.

**Table 2: Some Typical Commands** 

Command	A Number	Comments
\author \table	100 300	Author For tables
\table*	400	For wider tables



Figure 1: 1907 Franklin Model D roadster. Photograph by Harris & Ewing, Inc. [Public domain], via Wikimedia Commons. (https://goo.gl/VLCRBB).

## Citations and Bibliographies

The use of *BibTeX* for the preparation and formatting of one's references is strongly recommended. Authors' names should be complete — use full first names ("Donald E. Knuth") not initials ("D. E. Knuth") — and the salient identifying features of a reference should be included: title, year, volume, number, pages, article DOI, etc.

The bibliography is included in your source document with these two commands, placed just before the \end{document} command:

\bibliographystyle{ACM-Reference-Format}
\bibliography{bibfile}

where "bibfile" is the name, without the ".bib" suffix, of the BibTeX file.

Some examples. A paginated journal article [2], an enumerated journal article [7], a reference to an entire issue [6], a monograph (whole book) [19], a monograph/whole book in a series (see 2a in spec. document) [13], a divisible-book such as an anthology or compilation [9] followed by the same example, however we only output the series if the volume number is given [10] (so Editor00a's series should NOT be present since it has no vol. no.), a chapter in a divisible book [30], a chapter in a divisible book in a series [8], a multi-volume work as book [18], an article in a proceedings (of a conference, symposium, workshop for example) (paginated proceedings article) [3], a proceedings article with all possible elements [29], an example of an enumerated proceedings article [11],

an informally published work [12], a doctoral dissertation [5], a master's thesis: [4], an online document / world wide web resource [1, 24, 31], a video game (Case 1) [23] and (Case 2) [22] and [21] and (Case 3) a patent [28], work accepted for publication [25], 'YYYYYb'-test for prolific author [26] and [27]. Other cites might contain 'duplicate' DOI and URLs (some SIAM articles) [17]. Boris / Barbara Beeton: multi-volume works as books [15] and [14]. A couple of citations with DOIs: [16, 17]. Online citations: [31–33].

# **Appendices**

If your work needs an appendix, add it before the "\end{document}" command at the conclusion of your source document.

Start the appendix with the "appendix" command:

\appendix

and note that in the appendix, sections are lettered, not numbered. This document has two appendices, demonstrating the section and subsection identification method.

## REFERENCES

- Rafal Ablamowicz and Bertfried Fauser. 2007. CLIFFORD: a Maple 11 Package for Clifford Algebra Computations, version 11. Retrieved February 28, 2008 from http://math.tntech.edu/rafal/cliff11/index.html
- [2] Patricia S. Abril and Robert Plant. 2007. The patent holder's dilemma: Buy, sell, or troll? Commun. ACM 50, 1 (Jan. 2007), 36–44. https://doi.org/10.1145/1188913. 1188915
- [3] Sten Andler. 1979. Predicate Path expressions. In Proceedings of the 6th. ACM SIGACT-SIGPLAN symposium on Principles of Programming Languages (POPL '79). ACM Press, New York, NY, 226–236. https://doi.org/10.1145/567752.567774
- [4] David A. Anisi. 2003. Optimal Motion Control of a Ground Vehicle. Master's thesis. Royal Institute of Technology (KTH), Stockholm, Sweden.
- [5] Kenneth L. Clarkson. 1985. Algorithms for Closest-Point Problems (Computational Geometry). Ph.D. Dissertation. Stanford University, Palo Alto, CA. UMI Order Number: AAT 8506171.
- [6] Jacques Cohen (Ed.). 1996. Special issue: Digital Libraries. Commun. ACM 39, 11 (Nov. 1996).
- [7] Sarah Cohen, Werner Nutt, and Yehoshua Sagic. 2007. Deciding equivalances among conjunctive aggregate queries. J. ACM 54, 2, Article 5 (April 2007), 50 pages. https://doi.org/10.1145/1219092.1219093
- [8] Bruce P. Douglass, David Harel, and Mark B. Trakhtenbrot. 1998. Statecarts in use: structured analysis and object-orientation. In Lectures on Embedded Systems, Grzegorz Rozenberg and Frits W. Vaandrager (Eds.). Lecture Notes in Computer Science, Vol. 1494. Springer-Verlag, London, 368–394. https://doi.org/10.1007/ 3-540-65193-4\_29
- [9] Ian Editor (Ed.). 2007. The title of book one (1st. ed.). The name of the series one, Vol. 9. University of Chicago Press, Chicago. https://doi.org/10.1007/ 3-540-09237-4
- [10] Ian Editor (Ed.). 2008. The title of book two (2nd. ed.). University of Chicago Press, Chicago, Chapter 100. https://doi.org/10.1007/3-540-09237-4
- [11] Matthew Van Gundy, Davide Balzarotti, and Giovanni Vigna. 2007. Catch me, if you can: Evading network signatures with web-based polymorphic worms. In Proceedings of the first USENIX workshop on Offensive Technologies (WOOT '07). USENIX Association, Berkley, CA, Article 7, 9 pages.
- [12] David Harel. 1978. LOGICS of Programs: AXIOMATICS and DESCRIPTIVE POWER. MIT Research Lab Technical Report TR-200. Massachusetts Institute of Technology, Cambridge, MA.
- [13] David Harel. 1979. First-Order Dynamic Logic. Lecture Notes in Computer Science, Vol. 68. Springer-Verlag, New York, NY. https://doi.org/10.1007/3-540-09237-4

- [14] Lars Hörmander. 1985. The analysis of linear partial differential operators. III. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 275. Springer-Verlag, Berlin, Germany. viii+525 pages. Pseudodifferential operators.
- [15] Lars Hörmander. 1985. The analysis of linear partial differential operators. IV. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 275. Springer-Verlag, Berlin, Germany. vii+352 pages. Fourier integral operators.
- [16] IEEE 2004. IEEE TCSC Executive Committee. In Proceedings of the IEEE International Conference on Web Services (ICWS '04). IEEE Computer Society, Washington, DC, USA, 21–22. https://doi.org/10.1109/ICWS.2004.64
- [17] Markus Kirschmer and John Voight. 2010. Algorithmic Enumeration of Ideal Classes for Quaternion Orders. SIAM J. Comput. 39, 5 (Jan. 2010), 1714–1747. https://doi.org/10.1137/080734467
- [18] Donald E. Knuth. 1997. The Art of Computer Programming, Vol. 1: Fundamental Algorithms (3rd. ed.). Addison Wesley Longman Publishing Co., Inc.
- [19] David Kosiur. 2001. Understanding Policy-Based Networking (2nd. ed.). Wiley, New York, NY.
- [20] Leslie Lamport. 1986. Lamport. Addison-Wesley, Reading, MA.
- [21] Newton Lee. 2005. Interview with Bill Kinder: January 13, 2005. Video. Comput. Entertain. 3, 1, Article 4 (Jan.-March 2005). https://doi.org/10.1145/1057270. 1057278
- [22] Dave Novak. 2003. Solder man. Video. In ACM SIGGRAPH 2003 Video Review on Animation theater Program: Part I - Vol. 145 (July 27–27, 2003). ACM Press, New York, NY, 4. https://doi.org/99.9999/woot07-S422
- [23] Barack Obama. 2008. A more perfect union. Video. Retrieved March 21, 2008 from http://video.google.com/videoplay?docid=6528042696351994555
- [24] Poker-Edge.Com. 2006. Stats and Analysis. Retrieved June 7, 2006 from http://www.poker-edge.com/stats.php
- [25] Bernard Rous. 2008. The Enabling of Digital Libraries. Digital Libraries 12, 3, Article 5 (July 2008). To appear.
- [26] Mehdi Saeedi, Morteza Saĥeb Zamani, and Mehdi Sedighi. 2010. A library-based synthesis methodology for reversible logic. Microelectron. J. 41, 4 (April 2010), 185–104
- [27] Mehdi Saeedi, Morteza Saheb Zamani, Mehdi Sedighi, and Zahra Sasanian. 2010. Synthesis of Reversible Circuit Using Cycle-Based Approach. J. Emerg. Technol. Comput. Syst. 6, 4 (Dec. 2010).
- [28] Joseph Scientist. 2009. The fountain of youth. Patent No. 12345, Filed July 1st., 2008, Issued Aug. 9th., 2009.
- [29] Stan W. Smith. 2010. An experiment in bibliographic mark-up: Parsing metadata for XML export. In Proceedings of the 3rd. annual workshop on Librarians and Computers (LAC '10, Vol. 3), Reginald N. Smythe and Alexander Noble (Eds.). Paparazzi Press, Milan Italy, 422–431. https://doi.org/99.9999/woot07-S422
- [30] Asad Z. Spector. 1990. Achieving application requirements. In Distributed Systems (2nd. ed.), Sape Mullender (Ed.). ACM Press, New York, NY, 19–33. https://doi.org/10.1145/90417.90738
- [31] Harry Thornburg. 2001. Introduction to Bayesian Statistics. Retrieved March 2, 2005 from http://ccrma.stanford.edu/~jos/bayes/bayes.html
- [32] TUG 2017. Institutional members of the TeX Users Group. Retrieved May 27, 2017 from http://wwtug.org/instmem.html
- [33] Boris Veytsman. [n.d.]. acmart—Class for typesetting publications of ACM. Retrieved May 27, 2017 from http://www.ctan.org/pkg/acmart

# A RESEARCH METHODS (REMOVE IF NOT USED)

### A.1 Part One

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi malesuada, quam in pulvinar varius, metus nunc fermentum urna, id sollicitudin purus odio sit amet enim. Aliquam ullamcorper eu ipsum vel mollis. Curabitur quis dictum nisl. Phasellus vel semper risus, et lacinia dolor. Integer ultricies commodo sem nec semper.

### A.2 Part Two

Etiam commodo feugiat nisl pulvinar pellentesque. Etiam auctor sodales ligula, non varius nibh pulvinar semper. Suspendisse nec lectus non ipsum convallis congue hendrerit vitae sapien. Donec at laoreet eros. Vivamus non purus placerat, scelerisque diam eu, cursus ante. Etiam aliquam tortor auctor efficitur mattis.

#### B ONLINE RESOURCES

Nam id fermentum dui. Suspendisse sagittis tortor a nulla mollis, in pulvinar ex pretium. Sed interdum orci quis metus euismod, et sagittis enim maximus. Vestibulum gravida massa ut felis suscipit congue. Quisque mattis elit a risus ultrices commodo venenatis eget dui. Etiam sagittis eleifend elementum.

Nam interdum magna at lectus dignissim, ac dignissim lorem rhoncus. Maecenas eu arcu ac neque placerat aliquam. Nunc pulvinar massa et mattis lacinia.