

Sample Paper for CLIB 2020

First Author

Affiliation

Affiliation

email@domain.com

Second Author

Affiliation

Affiliation

email@domain.com

Abstract

This document contains the specifications for preparing a camera-ready manuscript for the proceedings of CLIB 2020. Authors should adhere to the directions described herein, both for papers submitted for review and for final versions of accepted papers. This document is an example of the CLIB 2020 template.

Type the abstract between the email address(es) and the main body of the text. The width of the abstract text should be smaller than the main body of the paper by 2 cm on each side. Centre the word **Abstract** in a 12 pt bold font above the body of the abstract. The abstract should be a concise summary of the general thesis and conclusions of the paper. It should be no longer than 250 words.

Keywords: at least 3 keywords

1. Introduction

This document has been compiled from several scientific article styles, including the guidelines for the COLING 2016 and ACL 2016 proceedings. Authors submitting a paper to CLIB 2020 should adhere to these instructions when preparing their papers, as should authors of accepted papers when submitting the final versions of their papers. The only acceptable format is the Portable Document Format (PDF).

The proceedings are designed for A4 paper format. Note that page numbers and proceedings footer are added by the organisers. **Please do not number the pages and do not place footers or headers.**

Any additional specifications will be made available on the Conference website:

<http://dcl.bas.bg/clib/>

Please check this website regularly for updates.

2. General Instructions

The electronic manuscript must be submitted in Adobe's Portable Document Format (PDF). PDF files are usually produced from \LaTeX using the *pdflatex* command. If your version of \LaTeX produces Postscript files, or your document requires processing into ps first (e.g. uses ps-related packages), you can convert these into PDF using *ps2pdf* or *dvipdf*. On Windows, you can also use Adobe Distiller to generate PDF.

Please make sure that your PDF file includes all the necessary fonts (especially tree diagrams, symbols, and fonts with special characters).

2.1. Templates

We strongly recommend that you prepare your PDF files using \LaTeX , preferably the $\text{\LaTeX}2\text{e}$ version with the official CLIB 2020 style file (clib2020.sty). The style file, as well as the document clib2020.pdf and its \LaTeX source code (clib2020.tex) are available in the Instructions for authors section of the website <http://dcl.bas.bg/clib/>.

2.2. MS Word / OpenOffice / LibreOffice

You can alternatively use a text editor, such as Microsoft Word or LibreOffice, to produce your PDF file. In this case, you should use the Word/LibreOffice template files (clib2020.dot and clib2020.ott, respectively) on the CLIB 2020 website. Make sure to anonymise your source file so that the pdf produced does not retain your identity. This can be done by removing any personal information from your source document properties.

3. Layout

This section contains guidelines for the paper layout.

3.1. General formatting

Make sure to specify the **A4 format** (21 cm x 29.7 cm) when formatting the paper. The exact dimensions for a page on A4 paper are:

- Left and right margins: 2.5 cm
- Top margin: 2.5 cm
- Bottom margin: 2.5 cm
- Width: 16.0 cm
- Height: 24.7 cm

Manuscripts must be in single-column format.

3.2. Fonts

Authors should use the default L^AT_EX2e font types (**Computer Modern**) or **Times Roman** (using L^AT_EXpackage `times`). With the Microsoft Word template (clib2020.dot) it is recommended to use the **Times New Roman** font. The LibreOffice template (clib2020.ott) uses FreeSerif for Linux.

| Type of Text | Font Size | Style |
|---------------------|-----------|-------|
| paper title | 15 pt | bold |
| author names | 12 pt | bold |
| author affiliation | 12 pt | |
| the word “Abstract” | 12 pt | bold |
| section titles | 12 pt | bold |
| document text | 11 pt | |
| captions | 11 pt | |
| sub-captions | 11 pt | |
| abstract text | 11 pt | |
| bibliography | 11 pt | |
| footnotes | 9 pt | |

Table 1: Font guide

3.3. The First Page

The title, author(s)’ name(s), affiliation(s) and electronic mail address(es) must be centred across the top of the first page, as indicated on the title page of this document.

Title: The title should be centred at the top of the first page, in a 15 pt bold font. (For a complete guide to font sizes and styles, see Table 1). Use title-case capitalisation. Long titles should be typed on two lines without a blank line intervening. Put the title at 2.5 cm from the top of the page, followed by a blank line, then the author(s)’ name(s), and the affiliation(s) on the following line. Do not include the paper ID number assigned during the submission process.

Authors: In case of multiple authors, there are several possible layout formats depending on the number of authors. The different variants for formatting author(s) name(s), affiliation(s) and email address(es) are included as comments in the preamble of this document. If the affiliation of all or part of the authors is the same, you can collapse the respective Affiliation line(s).

Use full given names (middle initials are allowed). Only the first letter of surnames should be capitalised. The affiliation should contain the author's home institution and an electronic mail address. In case of same-domain emails of multiple authors, they can be collapsed as indicated `email1, email2, email3@domain.com`

The command `\maketitle` produces the title and authors layout with the predefined template options.

Make sure that the title, author(s) name(s) and affiliation(s) are completely identical to those entered in the electronic paper submission website.

3.4. Sections and subsections

Sections and subsections should be numbered automatically. Unnumbered (sub)sections can be produced using

```
\section*{title}
\subsection*{title}
```

but should generally be avoided.

3.4.1. Headings

Type and label section and subsection headings as indicated in the present document. Use numbered sections (Arabic numerals) in order to facilitate cross references. Number subsections with the section number and the subsection number separated by a dot, in Arabic numerals. Use sentence-case capitalisation for section headings.

3.4.2. Text

Begin typing the main body of the text immediately after the abstract, observing the single-column format as shown in the present document. Do not indent the first paragraph of each section or subsection. Indent all other new paragraphs (the default is set at 0.7 cm from the left).

3.4.3. Appendices

Appendices, if any, directly follow the text and the references to the exception of the cases when they themselves contain references (see below the section on **References and citing**). Letter them in sequence and provide an informative title: **Appendix A. Title of Appendix.**

3.5. Footnotes

Put footnotes at the bottom of the page and use 9 pt text and number them in sequence.¹ Footnotes should be separated from the text by a line.² These are standard in both \LaTeX and Word.

3.6. Numbered and unnumbered lists

Both numbered and unnumbered lists can be used.

For numbered lists use the `enumerate` environment, and for unnumbered lists – the `itemize` environment. These environments can be embedded in one another, if needed.

1. First item
2. Second item
 - First subitem

¹This is how a footnote should appear.

²Note the line separating the footnotes from the text.

- Second subitem
 - default subsubsubitem bullet
 - another default subsubsubitem bullet
 - + custom subsubsubitem bullet

A custom bullet can be used for an item, as follows:

```
\item[+] subsubsubitem
```

3.7. Tables and graphics

Place figures, tables, and other graphic representations in the paper near where they are first discussed, rather than at the end, if possible. Figures, tables, and graphics should be centred.

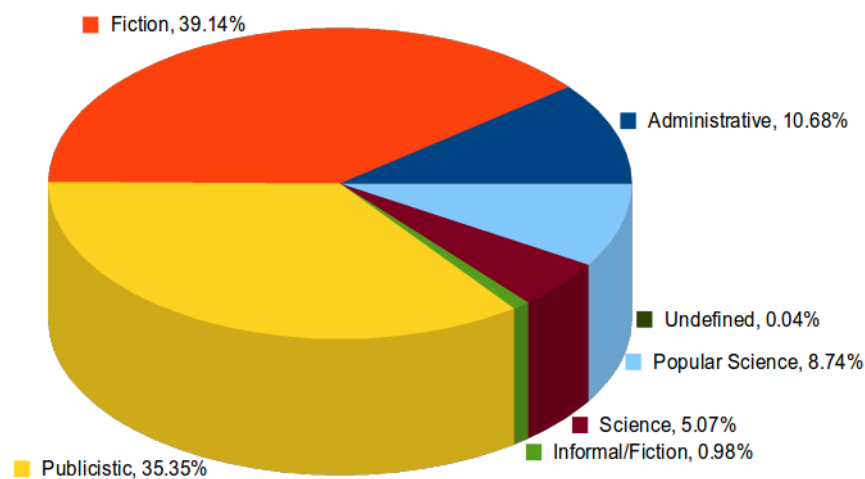


Figure 1: Caption of Figure 1

Captions: Provide a caption for every illustration or table; these should be numbered automatically in sequence in the following format: “Figure 1: Caption of the Figure” “Table 1: Caption of the Table”. Type the captions of the figures and tables below the body.

Narrow graphics together with the single-column format may lead to large empty spaces, see for example the wide margins on both sides of Figure 1. To save space, you can combine graphics and/or tables using either:

- sub-graphics with sub-captions, e.g. numbered (a), (b), (c), etc. if the graphics/tables are related; or
- minipage to place images/tables with independent captions side by side, as in Figure 2.

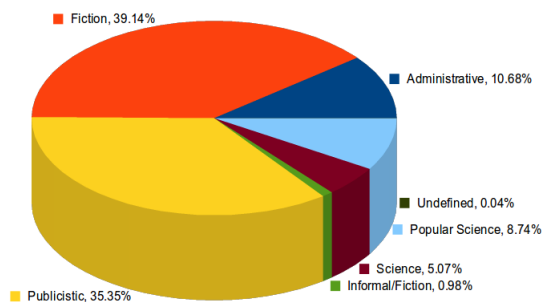


Figure 2: Caption of Figure 2

| Style | No texts | No words |
|-----------------|----------|----------|
| Administrative | 23111 | 357888 |
| Fiction | 2347 | 87401 |
| Journalistic | 102345 | 2319541 |
| Scientific | 981 | 76191 |
| Popular science | 1109 | 20081 |
| Popular | 308 | 10981 |
| Colloquial | 11 | 9087 |

Table 2: Caption of Table 2

Any of the \LaTeX packages `float`, `wrapfig`, `subfig`, can be used.

4. References and citing

The following rules apply for citations:

- Citations within the text `\cite{Fellbaum1998}` appear in parentheses as in (Fellbaum, 1998). Append lowercase letters ('a', 'b', ...) to the year in cases you cite more than one work published in the same year by the same author.
- Treat double authors `\cite{Koeva2011}` as in (Koeva and Genov, 2011).
- Write `\cite{Koeva2012}` as in (Koeva et al., 2012) when more than two authors are involved.
- Collapse multiple citations `\cite{Fellbaum1998,Burnard2005}` as in (Fellbaum, 1998; Burnard, 2005).
- For direct quotations, include a page number, separated from the year by a colon followed by a space, as indicated `\cite[35]{Fellbaum1998}` (Fellbaum, 1998: 35).
- Use a dash to cite page ranges, as in (Fellbaum, 1998: 35–37). If the cited author's name appears in the text itself, then put only the year (and possibly page number) in parentheses, as in "Fellbaum (1998: 33–50) argues that..."

`\newcite[33--50]{Fellbaum1998} argues...`

References should be compliant with the Harvard-style referencing format. The \LaTeX style file and the sample bibliography file provide examples of the format for different publications. Both full first names of authors and initials are acceptable.

Authors need to ensure consistency in the capitalisation of the titles included in the bibliography items. Different item types have different default behaviour with respect to capitalisation – titles of `@article` and `@inproceedings` use sentence-like capitalisation (first word only), while `@book` and `@techreport` use title-case capitalisation. To force a certain title behaviour, you can use double curly brackets, e.g.:

`title = {{The Bulgarian National Corpus}}`

List all references together in a separate unnumbered section under the heading **References**; place the section before any **Appendices**, unless they contain references. Arrange the references alphabetically by author's family name, followed by the year of publication. Provide as complete citations as possible.

It is recommended that you use a separate `bib` file for the bibliography to ensure consistency of style. In case you use the simplified format with `\bibitem`, it is your responsibility to observe the style guidelines given in this document.

Please refer to the bib file `clib2020.bib` for examples of how to describe references of: manuals (Christ and Schulze, 1994), technical reports (Briscoe, 2006), electronic sources (EAGLES, 1996), works in collections (Clear, 1992), etc.

5. Anonymous submission for blind review

As reviewing will be double-blind, the submitted version of the papers to be reviewed **should not** include the author(s)' name(s) and affiliation(s). Therefore, make sure to leave the `\author{}` in the preamble empty. Self-references that reveal the author's identity should be avoided. In such cases, instead of "In our previous work we showed..." use general citations such as:

"Koeva and Genov (2011) previously showed ..."

In addition, **do not include** any of the following when submitting your paper for review: acknowledgements, project names, grant numbers, and names or URLs of resources or tools that reveal author(s)' identity.

Papers that do not conform to the anonymity requirements may be rejected without review. Acknowledgements can be included in the final version of the paper for the proceedings.

6. Multilingual support and translation of non-English terms and examples

Authors are advised to supplement non-English terms, phrases or larger text fragments with appropriate transliterations and/or translations. It is highly recommended that words in non-Latin alphabets are transliterated instead of given in their original form.

Inline transliterations and/or translations should be formatted in the following order: (1) original form; (2) *transliteration* – in italic; and (3) "translation" – in double quotes, as in:

For the transliteration of Bulgarian alphabetic characters use the standard approved by the Government of the Republic of Bulgaria³. For other variants of the Cyrillic alphabet use an internationally accepted standard and provide references and link to it in the **References** section.

Authors can use the `babel` package for typesetting non-English text. The following lines need to be included in the preamble:

```
\usepackage[utf8]{inputenc}
\usepackage[english,bulgarian]{babel}
```

with the foreign words or text fragments included in the following way:

```
{\selectlanguage{bulgarian} text in another language}
```

7. Length of Submission

The maximum submission length is 8 pages (A4), plus up to two extra pages for references. Final versions should take into account reviewers comments.

Papers that do not conform to the specified length and formatting requirements may be rejected without review. If you cannot meet the above requirements for the production of your electronic submission, please contact the Organising Committee as soon as possible. Contact details can be found on the Conference website.

Acknowledgements

If you would like to include any acknowledgements, they should go in a separate **unnumbered** section immediately before the references. **Please do not include this section when submitting your paper for review in order to preserve anonymity.**

³<http://www.mrrb.government.bg/?controller=articles&id=403>

References

- Briscoe, T. (2006). *An Introduction to Tag Sequence Grammars and the RASP System Parser*. Technical report, University of Cambridge, Computer Laboratory Technical Report.
- Burnard, L. (2005). *Developing Linguistic Corpora: a Guide to Good Practice*, chapter Metadata for corpus work. Oxford: Oxbow Books. <http://ota.ahds.ac.uk/documents/creating/dlc/index.htm>.
- Christ, O. and Schulze, B. M. (1994). *The IMS Corpus Workbench: Corpus Query Processor (CQP) User's Manual*. University of Stuttgart, Germany.
- Clear, J. (1992). Corpus Sampling. In Leitner, G., Ed., *New Directions in English Language Corpora*. Berlin: Mouton de Gruyter.
- EAGLES. (1996). *EAGLES: Preliminary Recommendations on Corpus Typology. EAGLES Document EAG—TCWG—CTYP/P*. <http://www.ilc.cnr.it/EAGLES96/corpus typ/corpus typ.html>.
- Fellbaum, C., Ed. (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.
- Koeva, S. and Genov, A. (2011). Bulgarian Language Processing Chain. In *Proceedings of Integration of multilingual resources and tools in Web applications. Workshop in conjunction with GSCL 2011, University of Hamburg*.
- Koeva, S., Stoyanova, I., Leseva, S., Dimitrova, T., Dekova, R., and Tarpomanova, E. (2012). The Bulgarian National Corpus: Theory and Practice in Corpus Design. *Journal of Language Modelling*, 0(1):65–110.