bibcop: LaTeX Package
for Style Checking of `.bib` Files*

Yegor Bugayenko
`yegor256@gmail.com`

2023-01-10, 0.0.7

NB! You must run TEX processor with `--shell-escape` option and you must have Perl installed. This package doesn't work on Windows.

1 Introduction

This package scans a `.bib` file for style errors and emits warning messages if any issues are found (the package must be included before all other bibliography related packages):

```
\documentclass{article}
\usepackage{bibcop}
\begin{document}
\bibliographystyle{plain}
\bibliography{main}
\end{document}
```

Some warnings may be printed in the TEX log. Once the issues in the `main.bib` file are fixed, the warnings disappear.

Bibcop doesn't pay much attention the formatting of the `.bib` file. It doesn't emit warnings if a comma is missed after the last tag in a BIBTEX entry or the tags and their equation symbols are not aligned. Instead, Bibcop is focused on the content of the tags and their possible inconsistencies.

If the `.sty` file is used directly (without installing it into the TEX tree), the `bibcop.pl` file must also be placed next to it — it is the Perl script that does all the work of checking your `.bib` files. The `.sty` is just a simple wrapper around it.

The `\usepackage{bibcop}` must stay right after `\usepackage{biblatex}` (if BibLaTEX is used), otherwise there won't be any warnings from `bibcop`.

---

*The sources are in GitHub at [yegor256/bibcop](https://github.com/yegor256/bibcop)

2   Package Options

It's possible to configure the behavior of the package with the help of a few package options:

verbose   The verbose package option prints all debugging messages to the TEX log:

```
\usepackage[verbose]{bibcop}
```

script   The package depends on the bibcop.pl file, which is a Perl script that does all the work. This file is supposed to be located either in the current directory or in the texmf-dist/scripts/bibcop/ directory. Using the script option the location of the script may be explicitly defined (it is not recommended to use this option unless there is a special demand for it):

```
\usepackage[script=my-perl-script.pl]{bibcop}
```

3   The Rules

This is a more or less complete list of rules enforced on a .bib file:

types   Only @article, @book, @inproceedings, and @misc types of BIBTEX entries are allowed. Everything else, like @manual, @phdthesis, and many others are simply prohibited. The mentioned four should be enough for everything.

tags   There are pretty limited lists of allowed tags for each type of BIBTEX entry. The tags that are not in the list are prohibited to use.

doi   Every BIBTEX entry must have the doi tag, which is a unique Digital Object Identifier of the material that you reference. It seems to be a good practice, in order to avoid ambiguity, to always mention the DOI. It also recommended to use the doi package, in order to make all "doi" fields turned into hyper links.

caps   In the title, booktitle, and journal tags, all major words must be capitalized, as it is recommended by APA:

```
title = {A Preliminary Architecture for a Basic Data-Flow Processor}
```

Here, the leading "A" is capital because it opens the title. The word "for" and the article "a" are minor words, that's why they are in lower case. Both parts of the composite word "Data-Flow" are capitalized. Sometimes this rule may need to be violated, when there is custom capitalization, as it is done by the author of the paper. In order to do this, the words with custom capitalization must be wrapped in curled brackets, for example:

```
title = {Structured Programming {with} Go {To} Statements}
```

author   The author must contain a list of authors separated by "and". Each author must have two parts separated by a comma. The first part is the last name of the author, the second part is a list of their first names separated by a space, for example:

```
author = {Knuth, Donald E. and Duane, Bibby}
```

When the list of authors is too long, it's possible to say "and others":

```
author = {Knuth, Donald E. and others}
```

When first names are shortened to a single letter, it has to have a tailing dot. A specially pronounced author may be wrapped it in curled brackets, to make Bibcop ignore it:

```
author = {{Some author} and {I}}
```

shorts It is not allowed to shorten any words aside from the `author` tag, for example this is illegal:

```
journal = {J. Log. Comput.}
```

This must be replaced with the following:

```
journal = {Journal of Logic and Computation}
```

brackets The `title`, `booktitle`, and `journal` must be wrapped in double brackets, for example:

```
title = {{A Survey of Symbolic Execution Techniques}}
```

This is necessary in order to prevent down-casing of capitalized words, which is done by some bibliography styles.

year It is not allowed to mention the year inside the title of a conference or a journal, for example, this would be illegal:

```
booktitle = {{1994 IEEE International Conference on Computer Languages}},
```

The year should only be mentioned in the `year` tag, nowhere else. In the `year` tag only numbers are allowed:

```
year = {1994},
```

month The `month` may contain only a number:

```
month = {12},
```

volume The `volume` may contain only a number:

```
volume = {32},
```

number The `number` may contain only a number:

```
number = {132},
```

pages The `pages` may contain either a number or two numbers separated by a double dash:

```
pages = {145--163},
```

proceedings The `booktitle` in the `@inproceedings` entry must always start with "`Proceedings of the`", as in this example:

```
booktitle = {{Proceedings of the International
  Conference on Computer Languages}},
```

arXiv If the `archivePrefix` is present, the `eprint` and the `primaryClass` must also be present and must adhere to the formatting principles of [arXiv identifiers](#):

```
@misc{bugayenko2021,
  archivePrefix = {arXiv},
  eprint = {2111.13384},
  primaryClass = {cs.PL},
}
```

typography All tags in each BibTeX entry are checked for obeying the basic typography rules:

- No spaces are allowed in front of a comma, a semi-colon, a colon, a dot, a question mark, and an exclamation mark;

- A text may not end with a dot, a comma, a semi-colon, a colon, or a dash;

- A triple dash must be surrounded by spaces.

Unicode Any non-ASCII characters are prohibited in the entire `.bib` file.

4   Implementation

First, we include a few packages. We need [iexec](#) for executing Perl scripts:

```
1 \RequirePackage{iexec}
```

Then, we process package options:

```
2 \RequirePackage{pgfopts}
3 \pgfkeys{
4   /bibcop/.cd,
5   verbose/.store in=\bibcop@verbose,
6   script/.store in = \bibcop@script,
7 }
8 \ProcessPgfPackageOptions{/bibcop}
```

bibcop.pl Then, we find the Perl script:

```
9  \makeatletter
10 \ifdefined\bibcop@script\else
11   \IfFileExists{bibcop.pl}
12     {\gdef\bibcop@script{perl ./bibcop.pl}}
13     {\gdef\bibcop@script{bibcop}}
14 \fi
15 \message{bibcop: The Perl script is at '\bibcop@script'^^J}%
16 \makeatother
```

\bibliography Then, we re-define the \bibliography command:

```
17 \makeatletter
18 \ifdefined\bibliography
19   \let\bibcop@oldbibliography\bibliography
20   \renewcommand\bibliography[1]{%
21     \iexec{\bibcop@script\space
22       \ifdefined\bibcop@verbose--verbose\fi\space
23       --latex '#1.bib'}%
24     \message{bibcop: style checking finished^^J}%
25     \bibcop@oldbibliography{#1}%
26   }
27 \fi
28 \makeatother
```

\addbibresource Then, we re-define the \addbibresource command:

```
29 \makeatletter
30 \ifdefined\addbibresource
31   \let\bibcop@oldaddbibresource\addbibresource
32   \renewcommand\addbibresource[1]{%
33     \iexec{\bibcop@script\space
34       \ifdefined\bibcop@verbose--verbose\fi\space
35       --latex '#1'}%
36     \message{bibcop: style checking finished^^J}%
37     \bibcop@oldaddbibresource{#1}%
38   }
39 \fi
40 \makeatother
```

```
41 \endinput
```

Change History

Index

Numbers written in italic refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in roman refer to the code lines where the entry is used.