



bibcop: L^AT_EX Package for Style Checking of .bib Files*

Yegor Bugayenko
yegor256@gmail.com

2022-12-15, 0.0.4

NB! You must run T_EX 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 T_EX log. Once the issues in the main.bib file are fixed, the warnings disappear.

If the .sty file is used directly (without installing it into the T_EX 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 Bib_{La}T_EX is used), otherwise there won't be any warnings from 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 T_EX log:

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

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

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.

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 }  
7 \ProcessPgfPackageOptions{/bibcop}
```

bibcop.pl Then, we copy the Perl script using \VerbatimCopy from [verbatimcopy](#):

```
8 \RequirePackage{verbatimcopy}
9 \VerbatimCopy{bibcop.pl}{bibcop.tmp.pl}
10 \message{bibcop: File with Perl script 'bibcop.pl' copied^^J}%
```

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

```
11 \makeatletter
12 \ifdefined\bibliography
13 \let\bibcop@oldbibliography\bibliography
14 \renewcommand\bibliography[1]{%
15 \iexec{perl ". /bibcop.tmp.pl" \space
16 \ifdefined\bibcop@verbose--verbose\fi \space
17 --latex '#1.bib'}%
18 \message{bibcop: style checking finished^^J}%
19 \bibcop@oldbibliography{#1}%
20 }
21 \fi
22 \makeatother
```

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

```
23 \makeatletter
24 \ifdefined\addbibresource
25 \let\bibcop@oldaddbibresource\addbibresource
26 \renewcommand\addbibresource[1]{%
27 \iexec{perl ". /bibcop.tmp.pl" \space
28 \ifdefined\bibcop@verbose--verbose\fi \space
29 --latex '#1'}%
30 \message{bibcop: style checking finished^^J}%
31 \bibcop@oldaddbibresource{#1}%
32 }
33 \fi
34 \makeatother

35 \endinput
```

Change History

0.0.1		extensive Perl testing.	3
	General: First draft.		3
0.0.2		Package options introduced, the	
	General: Documentation extended,	<code>verbose</code> option enables detailed	
	more rules added.	logging inside the \TeX log.	3
0.0.4		The <code>-verbose</code> option introduced, to	
	General: Extra checks for the	enable debugging information only	
	typography, together with more	on demand.	3

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.

A		E	\ProcessPgfPackageOptions
\addbibresource <u>23</u>	\endinput 35
		 7
B		I	R
\bibcop.pl <u>8</u>	\iexec 15, 27
\bibcop@oldaddbibresource 25, 31	\ifdefined	. 12, 16, 24, 28
\bibcop@oldbibliography 13, 19		
\bibcop@verbose	5, 16, 28	M	
\bibliography <u>11</u>	\makeatletter 11, 23
		\makeatother 22, 34
		\message 10, 18, 30
		P	
		\pgfkeys 3
		S	
		\renewcommand 14, 26
		\RequirePackage	.. 1, 2, 8
		V	
		\VerbatimCopy 9