

The `censor` Package

Tools for Producing Redacted Documents

Steven B. Segletes
SSegletes@verizon.net

2022/04/21
v4.2

1 Introduction

The `censor` package allows a convenient redaction/censor capability to be employed, for those who need to protect restricted information, as well as for those who are forced to work in a less efficient environment when dealing with restricted information.

Let us assume you have a document for internal use, containing some restricted (*i.e.*, private, sensitive, proprietary or classified) information. To give a very short (fictitious) example:

RESTRICTED SOURCE: _____

The Liberty missile, with charge diameter (CD) of 80~mm,
revealed a penetration capability of 1.30 1.19, and
1.37~CD in three recent tests into armor steel.

RESTRICTED OUTPUT: _____

The Liberty missile, with charge diameter (CD) of 80 mm, revealed a
penetration capability of 1.30 1.19, and 1.37 CD in three recent tests into
armor steel.

Censor/redaction capabilities are desirable to sanitize the information. This would accomplish two things: allow for wider distribution of the output and/or allow a less sensitive output to be stored locally with fewer storage controls.

There are two modes in which the `censor` package can be used: (i) when the source (`.tex` file) remains restricted/secure; and (ii) when the source (`.tex` file) is public/unsecure.

2 Mode I: Restricted/Secure Source (.tex file)

In this mode, you create the source in a restricted/secure environment, but would like to produce output that can be more widely distributed, or stored in a less restricted location. With the addition of `\usepackage{censor}` to your document preamble, the `\censor` command becomes accessible to block out key identifiers:

CENSORED RESTRICTED SOURCE: _____

```
The \censor{Liberty} missile, with charge diameter (CD) of
\censor{80}~mm, revealed a penetration capability of 1.30
1.19, and 1.37~CD in three recent tests into armor steel.
```

CENSORED NO-LONGER-RESTRICTED OUTPUT: _____

The ██████ missile, with charge diameter (CD) of ██ mm, revealed a penetration capability of 1.30 1.19, and 1.37 CD in three recent tests into armor steel.

The censored version of the output is (presumably) less sensitive in its content, and may be stored in or distributed to a less sensitive environment (*e.g.*, as a hardcopy). The censored words are not part of the output document (hardcopy or electronic), even though the space allocated for the redaction is identical to the original word being redacted.

The document may also be printed out in its restricted (uncensored) form, merely with the addition of the `\StopCensoring` command at the beginning of the document:

UNCENSORED RESTRICTED SOURCE: _____

```
\StopCensoring
```

```
The \censor{Liberty} missile, with charge diameter (CD) of
\censor{80}~mm, revealed a penetration capability of 1.30,
1.19, and 1.37~CD in three recent tests into armor steel.
```

UNCENSORED RESTRICTED OUTPUT: _____

The Liberty missile, with charge diameter (CD) of 80 mm, revealed a penetration capability of 1.30, 1.19, and 1.37 CD in three recent tests into armor steel.

As of version 3.0, the ability to censor “boxed” objects like images and/or tabular environments is made possible with the `\censorbox` command. As an example, one could use the following source:

UNCENSORED RESTRICTED SOURCE: _____

The `\censor{Liberty}` missile, with charge diameter (CD) of `\censor{80}`~mm, revealed a penetration capability into armor steel, as detailed in Table 1.

```

\begin{table}[ht]
\begin{center}
\textbf{Table 1. \censor{Liberty} Missile Test Data}\
\censorbox{%
  \small\begin{tabular}{cc}
Standoff & Penetration \\
(CD) & (CD) \\
\hline
5.0 & 1.30 \\
6.0 & 1.19 \\
6.8 & 1.37
\end{tabular}%
}
\end{center}
\end{table}

```

in order to turn, what would otherwise be the following restricted output into a censored, unrestricted table.

UNCENSORED RESTRICTED OUTPUT: _____

The Liberty missile, with charge diameter (CD) of 80 mm, revealed a penetration capability into armor steel, as detailed in Table 1.

Table 1. Liberty Missile Test Data

Standoff (CD)	Penetration (CD)
5.0	1.30
6.0	1.19
6.8	1.37

CENSORED UNRESTRICTED OUTPUT: _____

The [REDACTED] missile, with charge diameter (CD) of [REDACTED] mm, revealed a penetration capability into armor steel, as detailed in Table 1.

Table 1. [REDACTED] Missile Test Data



As of version 2.0, there is provided a block-censor capability, for redacting larger blocks of text. This new command is `\blackout`, and is used in the form `\blackout{Block of text}`. This nice thing about this command is that it can stretch across line, paragraph, and page boundaries conveniently. For example,

UNCENSORED RESTRICTED SOURCE: _____

```
\blackout{
The Liberty missile, with charge diameter (CD) of 80~mm,
revealed a penetration capability of 1.30, 1.19, and 1.37~CD
in three recent tests into armor steel.}

New paragraph.}
```

CENSORED UNRESTRICTED OUTPUT: _____



Prior to V4.0, `\blackout` did not conveniently work across environment boundaries, such as `math` or `tabular`. To blackout the whole `tabular` environment, one could use the `\censorbox` command as has been shown. Presently, however, there is another option—to censor cell-by-cell using a shorthand notation.

The `\blackout` macro now allows uncensored content to be *escaped* from its argument by delimiting it between `|...|` bars (note: the closing escape bar must be in the same grouping level). Thus, the censoring directive `\blackout{This |is a| test}` renders as [REDACTED] is a [REDACTED].

This escaping of text becomes essential if one wishes to execute environments within a `\blackout` argument. Unless the environment name is escaped, it will break the command, since the environment name itself would be replaced with censoring commands. As of v4.2, tokens of catcode 4 (the `&` token) are intercepted and automatically passed through to the output stream, which can be of use in censoring `tabular` expressions. Here shows how a partial and full censor could be accomplished, with the use of appropriate text-escaping delimiters, instead of a `\censorbox`:

```

\blackout{
|\begin{tabular}{|c|c|}
\hline
Title 1 & Title 2 \\
\hline
a & |bbb\\
cc| & dd\\
\hline
\end{tabular}|
|\begin{tabular}{|c|c|}|
\hline
Title 1 & Title 2 \\
\hline
a & bbb \\
cc & dd\\
\hline
|\end{tabular}|
}

```

Title 1	Title 2	■■■■■	■■■■■
a	■■■	■	■■■
■	dd	■	■

For comparison, the uncensored tabulars are:

Title 1	Title 2	Title 1	Title 2
a	bbb	a	bbb
cc	dd	cc	dd

****v4.0 CHANGE**:** The restriction on `\blackout` *not* ending on glue is no longer applicable.

Finally, the argument to `\blackout` can, in limited situations, process macros; However, the macros are not expanded by default. See section 5 for details on what is possible with respect to the censoring of macro content with this package.

3 Mode II: Public/Unsecure Source (.tex file)

This mode is envisioned for users who must bear a large level of inconvenience to work in a restricted/secure environment (generally meaning in a location physically removed from one's desk). Its use is envisioned to allow a large percentage of a document to be created in an public/unsecure environment, with only the restricted details being completed in a restricted/secure environment.

After including the `\usepackage{censor}` command in the document preamble, the `\censor*` and `\censorbox*` commands becomes accessible. The way this author envisions its use is as follows:

CENSORED UNRESTRICTED SOURCE: _____

```
% KEYWORD IDENTIFIERS:
\def\Missile{\censor*{7}}
\def\Size{\censor*{2}}
\def\TableOne{\censorbox*[\small]{26}{5}{3.5}}
```

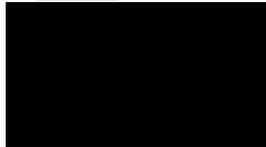
The `\Missile{}` missile, with charge diameter (CD) of `\Size` mm, revealed a penetration capability ranging from 1.19--1.37 CD in three recent tests into armor steel, as detailed in Table 1.

```
\begin{table}[ht]
\begin{center}
\textbf{Table 1. \Missile{ } Missile Test Data}
\end{center}
\end{table}
```

CENSORED UNRESTRICTED OUTPUT: _____

The ██████ missile, with charge diameter (CD) of █ mm, revealed a penetration capability ranging from 1.19–1.37 CD in three recent tests into armor steel, as detailed in Table 1.

Table 1. ██████ Missile Test Data



In this way, the source (.tex file) contains no restricted information, and may thus be prepared in an unrestricted environment. The argument to the `\censor*` command is a number representing the approximate width of the redaction (in ex's). In the case of the `\censorbox*` command, it has three mandatory arguments and one optional argument (note that the unstarred form of the `\censorbox` command has not three, but one, mandatory arguments). The mandatory arguments to `\censorbox*` are box width (in ex's), box height (in multiples of `\baselineskip`) and finally the depth below the baseline where the bottom of the box should be (in multiples of `\baselineskip`). The optional argument can include commands that you want in force for the `\censorbox` command, most typically fontsize commands, which will affect the size of the an ex and `\baselineskip`.

Because the redaction width is only approximate, it is possible that the censored and uncensored originals might have differing text justification. In the text, empty braces may be added after the keyword identifier macros so as to produce the proper interaction with the subsequent whitespace and punctuation.

Because the source (.tex) file contains no restricted information, the use of `\StopCensoring` cannot (in and of itself) produce the restricted document. Rather it will produce the following:

UNCENSORED UNRESTRICTED OUTPUT: _____

The _____ missile, with charge diameter (CD) of __ mm, revealed a penetration capability ranging from 1.19–1.37 CD in three recent tests into armor steel, as detailed in Table 1.

Table 1. _____ Missile Test Data

A rectangular box with a thin black border, representing a missing table. The box is empty and centered horizontally below the caption.

In order to produce the uncensored, restricted output, the unrestricted source file must be moved into the restricted/secure environment and be subject to a small amount of additional editing. Once in the restricted environment, the source (.tex file) may be edited such that the censored keyword identifiers are filled in with their restricted values and, in the process, changing the `\censor*` and `\censorbox*` commands to `\censor` and `\censorbox`, respectively. This and the addition of the `\StopCensoring` command to the file will produce the uncensored, restricted result:

UNCENSORED NO-LONGER-UNRESTRICTED SOURCE: _____

```
\StopCensoring
\def\Missile{\censor{Liberty}}
\def\Size{\censor{80}}
\def\TableOne{\censorbox{%
  \begin{tabular}{cc}
    Standoff & Penetration \\
    (CD) & (CD) \\
    \hline
    5.0 & 1.30 \\
    6.0 & 1.19 \\
    6.8 & 1.37 \\
  \end{tabular}%
}}
```

The {\Missile} missile, with charge diameter (CD) of {\Size}~mm, revealed a penetration capability ranging from 1.19--1.37~CD in three recent tests into armor steel, as detailed in Table 1.

```
\begin{table}[ht]
\begin{center}
\textbf{Table 1. {\Missile} Missile Test Data}\
{\TableOne}
\end{center}
\end{table}
```

UNCENSORED RESTRICTED OUTPUT: _____

The Liberty missile, with charge diameter (CD) of 80 mm, revealed a penetration capability ranging from 1.19–1.37 CD in three recent tests into armor steel, as detailed in Table 1.

Table 1. Liberty Missile Test Data

Standoff (CD)	Penetration (CD)
5.0	1.30
6.0	1.19
6.8	1.37

The only changes required of the document were at the very beginning, among the keyword identifiers. The primary text of the source document remained un-

altered. Note also that the censored, unrestricted output may also be obtained from the no-longer-unrestricted source by removing the `\StopCensoring` command at the beginning of the file.

It is not envisioned that the `\blackout` command, discussed in the prior section, is used in this mode, where the source file is in an unsecure environment.

4 Contiguous Blackout

The author received some negative feedback concerning the appearance resulting from the `\blackout` command. Generally, people did not like the fact that spaces were not blacked out. To this end, the `\xblackout` command is introduced in V3.10, which *cosmetically* blacks out the text more completely (note, however, that underneath the more complete blackout are still the word-length blackouts that otherwise compose `\blackout`).

As of v4.2 of this package, additional blacking-out provided by the `\xblackout` macro does not make use of any tunable parameters, but merely employs the `\fontdimen` parameters already provided by each font.

The blacking from `\xblackout` is intended to be less “choppy” than that of `\blackout`. However, it too has some drawbacks. Importantly, because spaces are now subject to blackout, any space that would otherwise be ignored by `\ignorespaces` or `\ignorespacesafterend` will not be ignored during the `\xblackout`. These will include bounding spaces in `tabular` cells as well as trailing spaces after the `\end` of environments.

With these caveats, it may still be found preferable to the `\blackout` macro, as shown in the example below.

UNCENSORED RESTRICTED SOURCE: _____

```
\xblackout{
The Liberty missile, with charge diameter (CD) of 80~mm,
revealed a penetration capability of 1.30, 1.19, and 1.37~CD
in recent testing here.  The tests were into armor steel.}

New paragraph} with open text.
```

UNCENSORED RESTRICTED OUTPUT: _____

The Liberty missile, with charge diameter (CD) of 80 mm, revealed a penetration capability of 1.30, 1.19, and 1.37 CD in recent testing here. The tests were into armor steel.

censoring `tabular` as well as `amsmath` expressions.)

Because macros are, likewise, passed through by default without censoring, this means the expression “the equation `\frac{E}{m}=c^2`,” will be rendered under `\blackout` as . There are several points that *must* be noted in a case such as this. First, the arguments to super/sub-scripts must be embraced, even if a single character in length. Secondly, censored math scripts will not be reduced to script size, but will be censored using the height and depth of the pre-defined censor rule (this can affect the vertical spacing of lines). Finally, customary math spacing associated with operators, punctuation, etc., may be lost, depending on the expression.

As of v4.2, `\blackout` will, by default, automatically detect and censor macros of math-mode Greek letters, such as `\alpha`, `\beta`, etc. If one, however, wishes to not waste compilation time searching for such instances, one may disable the feature with the switch `\censormathgreekfalse`.

7 Censor-Rule Marks

The censor (blacking) mark is a `\rule` which has a depth $-0.3ex$ (*i.e.*, below the baseline), and a total height of $2.1ex$, by default. The characteristics of that rule may be changed by setting the following two lengths.

```
\censorruledepth=length
\censorruleheight=length
```

When using `\blackout` and `\xblackout`, periods are now censored by default. If, however, you wish the prior approach of not censoring periods, you may redefine `\censordot` to something else, for example,

```
\renewcommand\censordot{.}
```

to make them visible in the censored document.

8 Usage Summary

The complete set of commands available to the `cancel` package, to bring about text redaction, are:

```
\cancel{text}
\cancel*{width multiplier}
```


- It is preferable not to apply `\censor` to whitespace, or text justification could be adversely affected. If one wishes to censor a multi-word phrase, such as “Little Bo Peep,” it is recommended to do it as follows:

```
\xblackout{Little Bo Peep}
```

If such a phrase is to be used repeatedly through a document, it is most convenient to place it as a macro:

```
\newcommand\CenName{\xblackout{Little Bo Peep}}
```

such that subsequent reference is done indirectly:

```
We examine the life of \CenName{} in this report,  
assured that \CenName{} line breaks.
```

We examine the life of ██████████ in this report, assured that ██████████
██████████ line breaks.

- Censoring can still affect text justification, because censored words are never hyphenated across line boundaries, whereas the original text may have been. For this reason, both `\blackout` and `\xblackout` operate locally in `\sloppy` mode, to minimize instances of overfull `\hboxes`. (Note however, that effects of `\sloppy` will take hold only for blacked-out text that closes out a paragraph.) If necessary, hyphenation points (`\-`) or `\allowbreak` may be added to words subject to `blackout` or `xblackout`, to facilitate manual hyphenation.
- Users have asked for color support in the package. I do not wish to hard-wire color into the package—however, if the `xcolor` package is loaded, the following preamble definition will force censoring to be in the selected color (in this example, red):

```
\def\censorcolor{red}  
\let\svcensorrule\censorrule  
\renewcommand\censorrule[1]{%  
  \textcolor{\censorcolor}{\svcensorrule{#1}}}
```

Let’s see if it works:

```
Is this blacked out text censored in red?
```



This technique can be used at the user’s discretion to provide color censoring. The value of `\censorcolor` can be redefined on the fly to change the color of the censoring.

9 Source Code

The source code for censor is included below:

```
% censor.sty
\def\censorversionnumber{4.2}
\def\censorversiondate{2022/04/21}
\ProvidesPackage{censor}
[\censorversiondate\ \censorversionnumber\
 Provides capability for redaction of sensitive information]
%
% This work may be distributed and/or modified under the
% conditions of the LaTeX Project Public License, either version 1.3
% of this license or (at your option) any later version.
% The latest version of this license is in
% http://www.latex-project.org/lppl.txt
% and version 1.3 or later is part of all distributions of LaTeX
% version 2005/12/01 or later.
%
% This work has the LPPL maintenance status 'maintained'.
%
% The Current Maintainer of this work is Steven B. Segletes.

\usepackage{pbox}

\newlength\censorruledepth
\newlength\censorruleheight

\censorruledepth=-0.3ex% -0.3ex DEFAULT
\censorruleheight=2.1ex% 2.1ex DEFAULT
\def\censordot{\censor{.}}% versus \def\censordot{.}%

\newcommand\censorrule[1]{\protect\rule[\censorruledepth]{#1}{\censorruleheight}}

\DeclareRobustCommand\censor{\@ifstar{\@cenlen}{\@cenword}}
\newcommand\@cenlen[1]{\censorrule{#1 ex}}
\newcommand\@cenword[1]{\censorrule{\widthofpbox{#1}}}

\newcommand\un@censor{\@ifstar{\un@cenlen}{\un@cenword}}
\newcommand\un@cenlen[1]{\protect\underline{\hspace{#1 ex}}}
\newcommand\un@cenword[1]{#1}

\newcommand\StopCensoring{%
\let\censor\un@censor%
\let\censorbox\un@censorbox%
\renewcommand\censpace{ }%
}
\newcommand\RestartCensoring{%
```

```

        \renewcommand\censor{\@ifstar{\@cenlen}{\@cenword}}%
        \renewcommand\censorbox{\@ifstar{\censor@dim}{\censor@box}}%
        \let\censpace\sv@censpace%
    }

%%%%

\usepackage{tokcycle}[2021/03/10]

\def\@dump#1{\if\relax\detokenize\expandafter{#1}\relax\else
  \addcytoks[1]{\expandafter\censor\expandafter{#1}}\fi\def#1{}}
\def\@append#1#2{\tc@defx#1{#1#2}}

\def\censpace{%
  \leaders\hrule height \dimexpr\censorruleheight+\censorruledepth\relax
  depth -\censorruledepth
  \hskip\fontdimen2\font plus \fontdimen3\font minus \fontdimen4\font
  \mbox{}}
\let\sv@censpace\censpace

\long\def\blackout#1{\protect\blackoutenv#1\endblackoutenv}
\long\def\xblackout#1{\protect\xblackoutenv#1\endxblackoutenv}

\newif\ifexpandarg

\xtokcycleenvironment\blackoutenv
  {\ifx.##1\@dump\censored@word\addcytoks[1]{\censordot}\else
    \ifnum\catcode'##1=3\relax\@dump\censored@word\addcytoks{##1}\else
    \ifnum\catcode'##1=4\relax\@dump\censored@word\addcytoks{##1}\else
    \ifnum\catcode'##1=7\relax\@dump\censored@word\addcytoks{##1}\else
    \ifnum\catcode'##1=8\relax\@dump\censored@word\addcytoks{##1}\else
    \@append\censored@word{##1}%
    \tcpeek\@next\ifx\@next\@tcEscapeptr\@dump\censored@word\fi
    \fi\fi\fi\fi\fi}
  {\tctestifcon\ifexpandarg{\expandafter\processtoks\expandafter
    {\expanded{##1}}\@dump\censored@word}{\groupedcytoks{\processtoks{##1}}%
    \@dump\censored@word}}\expandargfalse}
  {\tctestifx{~##1}{\@append\censored@word{##1}}%
    {\tctestifx{\expanded##1}{\@dump\censored@word\expandargtrue}%
    {\test@chars{##1}\if@char\if@mathgreek
    \@append\censored@word{\ensuremath{##1}}\else
    \@append\censored@word{##1}\fi\else
    \test@accents{##1}\if@accent\@append\censored@word{##1}\tcpop\@tc@popped
    \expandafter\@append\expandafter\censored@word\expandafter{\@tc@popped}%
    \else\@dump\censored@word\addcytoks{##1}\fi\fi}}
  {\@dump\censored@word\addcytoks{##1}}
  {\stripgroupingtrue\def\censored@word{ }\sloppy}
  {\@dump\censored@word}

```

```

\xtokcycleenvironment\xblackoutenv
  {\ifx.##1\@dump\censored@word\addcytoks[1]{\censordot}\else
    \ifnum\catcode'##1=3\relax\@dump\censored@word\addcytoks{##1}\else
    \ifnum\catcode'##1=4\relax\@dump\censored@word\addcytoks{##1}\else
    \ifnum\catcode'##1=7\relax\@dump\censored@word\addcytoks{##1}\else
    \ifnum\catcode'##1=8\relax\@dump\censored@word\addcytoks{##1}\else
    \@append\censored@word{##1}%
    \tcpeek\@next\ifx\@next\@tcEscapeptr\@dump\censored@word\fi
    \fi\fi\fi\fi\fi}
  {\tctestifcon\ifexpandarg{\expandafter\processtoks\expandafter
    {\expanded{##1}}\@dump\censored@word}{\groupedcytoks{\processtoks{##1}}%
    \@dump\censored@word}}\expandargfalse}
  {\tctestifx{~##1}{\@append\censored@word{##1}}}%
  {\tctestifx{\expanded##1}{\@dump\censored@word\expandargtrue}}%
  {\test@chars{##1}\if@char\if@mathgreek
    \@append\censored@word{\ensuremath{##1}}\else
    \@append\censored@word{##1}\fi\else
    \test@accents{##1}\if@accent\@append\censored@word{##1}\tcpop\@c@popped
    \expandafter\@append\expandafter\censored@word\expandafter{\@c@popped}%
    \else\@dump\censored@word\addcytoks{##1}\fi\fi}}
  {\@dump\censored@word\addcytoks{\ccenspace}}
  {\stripgroupingtrue\def\censored@word{ }\sloppy}
  {\@dump\censored@word}

\newcommand\ccenspace{\ifhmode\ifmmode\else\censpace\fi\fi}

\newif\if@char
\newif\if@mathgreek
\newif\if@accent

\newcommand\test@chars[1]{\@charfalse\@mathgreekfalse
  \ifx\ $#1\@chartrue\else
  \ifx\&#1\@chartrue\else
  \ifx\##1\@chartrue\else
  \ifx\%#1\@chartrue\else
  \ifx\_#1\@chartrue\else
  \ifx\o#1\@chartrue\else
  \ifx\O#1\@chartrue\else
  \ifx\oe#1\@chartrue\else
  \ifx\OE#1\@chartrue\else
  \ifx\aa#1\@chartrue\else
  \ifx\AA#1\@chartrue\else
  \ifx\ae#1\@chartrue\else
  \ifx\AE#1\@chartrue\else
  \ifx\l#1\@chartrue\else
  \ifx\L#1\@chartrue\else
  \ifcensormathgreek\test@mathgreek{##1}\if@mathgreek\@chartrue\fi\fi%
  \fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi\fi
}

```


- 4.1 - Added automatic support for censoring character macros such as `\$`, etc. and character accents such as `\'`, etc. See `\test@chars` and `\test@accents` for details.
 - Fixed bug in which `\expanded` arguments remained inside a TeX group
- 4.2 - Extended capability (via switch) to autocensor Greek math (see `\test@mathgreek`).
 - In `(x)blackout` macros/environments, auto-intercepts catcode 3,4,7 and 8 characters and passes them through unchanged, to help facilitate general censoring of math and tabular content.
 - Switched `\censpace`, used in `\xblackout`, to a leader rather than a lapped rule, which allows automatically for glue associated with `\fontdim`.
 - Made `\censor` and `\censorbox` robust, which will allow censoring to occur across, for example, section headers, tocs, etc.
 - Revisited the documentation, bringing it up to date.