

The `censor` Package

Tools for Producing Redacted Documents

Steven B. Segletes
steven.b.segletes.civ@mail.mil

July 30, 2013
v3.21

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 more inefficient 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.

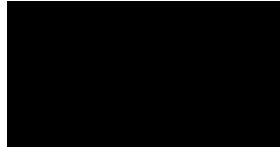
Table 1. Liberty Missile Test Data

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



Additionally, 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 and paragraph 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: _____



There is a caveat regarding the use of `\blackout`. The argument cannot end on so-called “glue,” such as a space or a carriage return. The following examples are, thus, bad forms of the `\blackout` command:

```
\blackout{Ends in space. Bad form }

\blackout{Ends in carriage return. Bad form
}
```

Secondly, periods are not censored (as of version 2.1). This is done because it is necessary to preserve the period-space combination that signifies end-of-sentence. In this way, L^AT_EX can preserve inter-word spacing in the redacted version that is identical to the unredacted version.

Thirdly, `\blackout` does not work across environment boundaries, such as math or tabular mode. You can blackout text in the table, cell by cell; But you cannot blackout the whole tabular environment with the `\blackout` command (instead, you should use the `\cancelbox` command for tabular environments).

Finally, the argument to `\blackout` can employ tokens; However, the tokens are expanded into one large blackout, regardless of whether the expanded token contains spaces. Thus

```
\blackout{\today}
```

becomes ██████████, in spite of the fact that `\today`, given at the moment by July 30, 2013, contains spaces.

The author received some negative feedback concerning the appearance resulting from a `\blackout` command. Generally, people did not like the fact that spaces were not blacked out, whether it was a concern over security or merely cosmetic. To this end, the `\xblackout` command is introduced in V3.10, which blacks out the text more completely, while simultaneously preserving the layout identical to that of the uncensored document.

The additional blacking-out of the `\xblackout` macro is controlled by five resettable lengths which lap additional blacking atop the result of `\blackout`. By default, these lengths are defined as:

```
\periodrlap = 1.6ex  
\afterperiodlap = 1.2ex  
\lletterlap = 0.55ex  
\rletterlap = 0.55ex  
\afterspacelap = 0.0ex
```

The length `\periodrlap` is the length of right-lapped blackout applied just before printing a period. Thus, this blacking is intended to cover the period not treated by `\blackout`. The length `\afterperiodlap` is a left-lapped blackout applied just prior to word following a period. Its intent is to back-cover the extra-wide spacing that follows a period. The length `\lletterlap` is the left-lapped blacking applied before a character. It will help cover inter-word space if `\afterspacelap` is not employed. The length `\rletterlap` is the right-lapped blacking applied after a character, which as with `\lletterlap`, can help cover inter-word spacing if `\afterspacelap` is not employed. Finally, there is `\afterspacelap` which, if employed, is a left-lapped blacking applied before the character that follows a space.

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}}
```

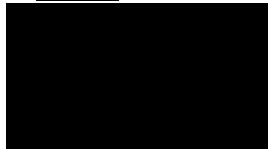
The `{\Missile}` missile, with charge diameter (CD) of `{\Size}`~mm, revealed a penetration capability ranging from 1.19--1.37~CD in three recent tests into armor steel, as detailed in Table 1.

```
\begin{table}[ht]
\begin{center}
\textbf{Table 1. {\Missile} Missile Test Data}\\
{\TableOne}
\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, the curly braces are placed around the keyword identifiers so as to produce the proper interaction with the surrounding 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



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` or `\xblackout` commands, discussed in the prior section, are used in this mode, where the source file is in an unsecure environment.

4 Censor Marks

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

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

In addition, it was mentioned that periods are not censored, so as to preserve the ability of L^AT_EX to produce censored output of *identical* spacing to the uncensored output. If, however, the preservation of space is not essential, and the user does not wish to have periods appear in the censored output, the definition `\censordot`, by default

```
\def\censordot{.}
```

can be redefined as a blank space, as a null character, or something else, at the users' discretion.

5 Summary & Miscellany

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

```
\censor{text}
\censor*{width mult.}
\censorbox[pre-commands]{object box}
\censorbox*[pre-commands]{width mult.}{height mult.}{depth mult.}
\blackout{extended text passage}
\xblackout{extended text passage}
\bpar (retained for backward compatibility only)
\StopCensoring
\RestartCensoring
```

```

\censorruledepth=length
\censorruleheight=length
\def\censordot{. or character}

```

The star (*) version of the commands is envisioned when the source document is being created in an unsecure environment, whereas, the unstarred version is when the document source may be created in a secure environment. In the star (*) commands, width multipliers are given in ex's, whereas height and depth multipliers are given in multiples of `\baselineskip`. The depth indicates the distance below the baseline where the bottom of a boxed object is to be placed (this command has a more direct effect for inline use, whereas use within environments is often overridden by the environment). The use of pre-commands will typically be fontsize commands, so that measurements of ex's and `\baselineskip` are done in a relevant fontsize.

The `\bpar` command is no longer needed, as of V3.10. It is now made equivalent to `\par`.

Censoring may be dynamically turned off and on in a document with the use of the `\StopCensoring` and `\RestartCensoring` commands, respectively. The default is censoring 'on.'

I have found that, in certain cases (for example, with captions created using the `boxhandler` package), the `cancelbox` command needs to be protected by way of `\protect\cancelbox{...}`.

It is preferable not to apply `\cancel` 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:

```
\blackout{Little Bo Peep}
```

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

```
\def\Name{\blackout{Little Bo Peep}}
```

such that subsequent reference is done indirectly:

```
We examine the life of {\Name} in this report
```

(Even so, it should be noted that censoring can still affect text justification, because censored words are never hyphenated across line boundaries, whereas the original text may have been.)

The source code for `cancel` is so short as to be included below:

```

% cancel.sty
\def\cancelversionnumber{3.21}
\ProvidesPackage{cancel}
[2013/06/24 \cancelversionnumber
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.

% VERSION:
% 1.00 - Initial release
% 2.00 - Added \blackout
% 2.10 - Allowed \blackout to cross paragraph boundaries with use of
%       \bpar. Stopped censoring periods, in order to preserve
%       end-of-sentence spacing, which differs from inter-word spacing.
% 3.00 - \censorbox introduced to handle figures, tables, etc.
% 3.10 - Made \blackout work with \par in argument. Introduced
%       \xblackout
% 3.20 - Specify depth/height of censor rule. Introduced \def\censordot{}
% 3.21 - Fixed bug regarding \xblackout rules remaining after a
%       \StopCensoring
\usepackage{pbox}
\usepackage{ifnextok}

\newlength\censorruledepth
\newlength\censorruleheight

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

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

\newcommand\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%
    \let\xblackout\blackout%
}
\newcommand\RestartCensoring{%
    \renewcommand\censor{\@ifstar{\@cenlen}{\@cenword}}%
    \renewcommand\censorbox{\@ifstar{\censor@dim}{\censor@box}}%
    \let\xblackout\sv\xblackout%
}

\let\sv@tilde~

\def\stringend{$}

\long\def\blackout#1{\def~{-}\censor@Block#1\stringend\let~\sv@tilde}
\long\def\censor@Block{\IfNextToken\stringend{\@gobble}%
    {\IfNextToken\@sptoken{ \bl@t{\censor@Block}}%
    {\bl@t{\censor@Block}}}}

% V2.00 DEFINITION:
% \def\bl@t#1#2{\censor{#2}#1}

% V2.10 DEFINITION (MADE \long IN V3.1):
\long\def\bl@t#1#2{\if\bpar#2\par\else\if.#2\censordot\else\censor{#2}\fi\fi#1}

%\def\bpar{ } %V3.00 DEFINITION
\let\bpar\par %AS OF V3.1, CAN HANDLE \par

\long\def\xblackout#1{\rule{0ex}{0ex}%
    \def~{-}%
    \def\@justpar{F}%
    \def\@justperiod{F}%
    \def\@justspace{F}%
    \xcensor@Block#1\stringend%
    \let~\sv@tilde%
}

\let\sv@xblackout\xblackout

\long\def\xcensor@Block{\IfNextToken\stringend{\@gobble}%
    {\IfNextToken\@sptoken{ \def\@justspace{T}\xbl@t{\xcensor@Block}}%
    {\xbl@t{\xcensor@Block}}}}

\newlength\periodrlap\setlength\periodrlap{1.6ex}
\newlength\afterperiodlap\setlength\afterperiodlap{1.2ex}

```

```

\newlength\lletterlap\setlength\lletterlap{0.55ex}
\newlength\rletterlap\setlength\rletterlap{0.55ex}
\newlength\afterspacelap\setlength\afterspacelap{0.0ex}

\def\@periodrlap{\rlap{\censorrule{\periodrlap}}}
\def\@afterperiodlap{\llap{\censorrule{\afterperiodlap}}}
\def\@lletterlap{\llap{\censorrule{\lletterlap}}}
\def\@rletterlap{\rlap{\censorrule{\rletterlap}}}
\def\@afterspacelap{\llap{\censorrule{\afterspacelap}}}

\long\def\xbl@t#1#2{%
  \if\par#2%
    \par\def\@justpar{T}%
  \else%
    \if T\@justspace%
      \rule{0ex}{1ex}\@afterspacelap%
    \fi%
    \if.#2%
      \def\@justperiod{T}%
      \@periodrlap%
      \censordot%
    \else%
      \if F\@justpar%
        \if T\@justperiod%
          \@afterperiodlap%
        \else%
          \@lletterlap%
        \fi%
      \fi%
      \cursor{#2}%
      \@rletterlap%
      \def\@justpar{F}%
      \def\@justperiod{F}%
      \def\@justspace{F}%
    \fi%
  \fi%
  #1%
}

\newcommand\cursorbox{\@ifstar{\cursor@dim}{\cursor@box}}
\newcommand\cursor@dim[4][]{\#1%
  \rule[-#4\baselineskip]{#2ex}{#3\baselineskip}}
\newcommand\cursor@box[2][]{\#1\setbox0\hbox{#2}%
  \rule[-\the\dp0]{\the\wd0}{\the\ht0+\the\dp0}}

\newcommand\un@cursorbox{\@ifstar{\un@cursor@dim}{\un@cursor@box}}

```

```
\newcommand\un@tensor@dim[4] [] []{#1%
    \fbox{\rule[-#4\baselineskip]{0ex}{#3\baselineskip}
    \rule{#2ex}{0ex}}}}
\newcommand\un@tensor@box[2] [] []{#1#2}
```

% NOTE: A \protect\censorbox{} MAY BE REQUIRED INSIDE SOME ENVIRONMENTS
\endinput