

Author Guidelines for *ICWSM* Proceedings*

Subtitle[†]

Nicolas Nicolov*
nicolas_nicolov@yahoo.com

Eytan Adar‡
eadar@cs.washington.edu

Natalie Glance◊
natalie.glance@buzzmetrics.com

Matthew Hurst◊
matthew.hurst@buzzmetrics.com

Mark Liberman‡
myl@cis.upenn.edu

Jim Martin△
martin@colorado.edu

Franco Salvetti♡
franco@powerset.com

*Umbria Inc., 1655 Walnut Str, Boulder, CO 80302, USA

‡Univ. of Washington, 101 Paul G. Allen Center, Box 352350, Seattle, WA 98195, USA

◊Nielsen BuzzMetrics, 5001 Baum Blvd, Suite 644, Pittsburgh, PA 15213, USA

‡Institute for Research in Cognitive Science, Univ. of Pennsylvania, 3401 Walnut St., Philadelphia, PA 19104, USA

△Dept. of Computer Science, Univ. of Colorado, Boulder, CO 80309, USA

♡Powerset Inc., 475 Brannan St., San Francisco, CA 94107, USA

Abstract

This document provides a L^AT_EX_{2 ϵ} sample (of the formatting) of a paper for the International Conference on Weblogs and Social Media (ICWSM). Authors need to compile it using L^AT_EX_{2 ϵ} and BibTeX. The formatting is adapted from the ACM's stylesheet `sig-alternate.cls`. The current document also uses some of the wording/examples/figures from the ACM's *Alternate ACM SIG Proceedings Paper in LaTeX Format*.

The developers of the ACM stylesheet have made an effort to include lots of “goodies”, such as a subtitle, footnotes on title, subtitle and authors, as well as in the text, and optional components (e.g., Appendices), not to mention examples of equations, theorems, tables and figures.

If the abstract is longer than half the textheight perhaps you are trying to say too much.

Keywords

Keywords are useful to the organizers to decide who should review a submission. Also the readers of the proceedings will find these helpful.

1. Introduction

This document is a sample of how to format a submission (and later, the camera ready copy) to the International Conference on Weblogs and Social Media (ICWSM). It is formatted using a popular scientific typesetting system (L^AT_EX) and is based on a template used by the Association for Computing Machinery (ACM). In the rest of the document we mention some of the features that are available to authors. More importantly, we also provide recommendations about the use of certain formatting features.

*Footnote on the title—to be avoided at all cost.

†A more detailed version of this technical report is available at www.uni.edu/~name/pubs/tr-007.pdf

We have two goals in writing these guidelines: (i) if you are not a L^AT_EX user we would like to show you some of the available functionality and make it easy for you to typeset your document in L^AT_EX; (ii) if you are a proficient L^AT_EX user we'd like the formatting to be as straightforward as possible. The main reason scientists use L^AT_EX is the quality of the final manuscript. People learning about the system often compare it to WYSIWYG (what you see is what you get) interfaces and find various aspects unintuitive (namely, you don't see how things appear). Actually there is a reason for this—L^AT_EX is a mark-up language and how things appear is dictated by templates. Other reasons for using L^AT_EX are: free availability and cross-platform compatibility.

This document is not a guide on good writing! It will not teach you *what to say* and in what order; it is neither on *how to say* it but on how to *typeset* it. Even worse, this document is not a short guide to the L^AT_EX system but rather a few comments on basic commands related to common structures of an ICWSM submission.

Here are some aspects of the formatting: fonts (Arial or Helvetica and Times Roman) in certain specified sizes (for instance, 9 point for body copy), a specified live area (18 × 23.5 cm [7" × 9.25"]) centered on the page, specified size of margins (2.54cm [1"] top and bottom and 1.9cm [.75"] left and right; specified column width (8.45cm [3.33"]) and gutter size (.083cm [.33"]). Don't worry—as an author you don't need to remember these; they will be taken care of by the template (stylesheet). The good news is, with only a handful of manual settings¹, the L^AT_EX document class file handles all of this for you.

The remainder of this document is concerned with showing, in the context of an “actual” document, the L^AT_EX commands specifically available for denoting the structure of a proceedings paper, rather than with giving rigorous descriptions or explanations of such commands.

¹ Two of these, the `\numberofauthors` and `\alignauthor` commands, you have already used; another, `\balancecolumns`, will be used in your very last run of L^AT_EX to ensure balanced column heights on the last page.

2. Anatomy of a paper

Let's start with the basic structure of a paper:

```
\documentclass{icwsm}
\begin{document}
\title{...}
\numberofauthors{1}
\author{
  \alignauthor John Doe\
  \affaddr{...}
  \email{...}
}
\date{}
\maketitle
\begin{abstract}
In this paper ...
\end{abstract}

\section{Introduction}
Weblogs are ...
\section{Formalism}
...
\section{Experiments}
...
\section{Conclusions}
...
\bibliographystyle{abbrv}
\bibliography{mybiblio}
\end{document}
```

Put this in a file `F.tex`; fill in the sections with your great thoughts (plain text) and you are mostly there. Then you run the commands:

```
latex F
bibtex F
latex F
latex F
dvips -Pcmz -f -t letter < F.dvi > F.ps
ps2pdf F.ps > F.pdf
```

This will ultimately produce the file `F.pdf`. Alternatively, you can modify the supplied `Makefile` to reflect the name of your submission and run `make`.

You probably would like to use more personal names than `F` or `icwsm`. Why? Because you don't want your submission to be overwritten by the submission of another person who decided to name their file using the same name.

2.1 The authors

At the beginning of the document there is a description of the authors. The stylesheet allows for up to three authors to be specified in vertical column format (you can have more authors—see further down). Here is an example with two authors:

```
\numberofauthors{2}
\author{
  \alignauthor John H. Watson
  \affaddr{Medical Corps} \
  \affaddr{Royal Army} \
  \affaddr{London, UK} \
  \email{watson@ramc.org}
  \alignauthor Sherlock Homes\titlenote{Has worked
    on cases with Dr.\,Watson.}\
  \affaddr{S.H.\ Consulting} \
  \affaddr{221b Baker Street} \
  \affaddr{London, England} \
  \email{sh@logic.co.uk}
}
```

It is also good to discuss among the co-authors the order in which the names will appear well in advance.

More than three authors

Now, if you have more than three authors it's probably better to use a different arrangement. In the source of this document we provide an example with seven authors where first the names and emails are in a vertical arrangement and the addresses are horizontally placed. The commmands for achieving this are just regular \LaTeX tables (`tabular` environment) we have also reduced the fontsize. We also commented out the `\numberofauthors{}` command:

```
%\numberofauthors{2}
\author{
  \begin{tabular}{cc}
Name Surname & Name2 Surname2 \
\texttt{email-1} & \texttt{email-2} \
\end{tabular} \
Company-1,... \
Company-2,... \
}
```

2.2 The body of the paper

Typically, the body of a paper is organized into a hierarchical structure, with numbered (or unnumbered) headings for sections, subsections, sub-subsections, and even smaller units. The command `\section` that precedes this paragraph is part of such a hierarchy.² \LaTeX handles the numbering and placement of these headings for you, when you use the appropriate heading commands around the titles of the headings. If you want a sub-subsection or smaller part to be unnumbered in your output, simply append an asterisk to the command name (e.g., `\subsection*{...}`). Examples of both numbered and unnumbered headings will appear throughout this document.

Use a blank line to represent a paragraph break.

3. Type changes and *special* characters

We have already seen several typeface changes in this sample. You can indicate *italicized* words or phrases in your text with the command `\textit{}`; **boldening** with the command `\textbf{}` and *typewriter-style* (for instance, for computer code) with `\texttt{}`. But remember, you do not have to indicate typestyle changes when such changes are part of the *structural* elements of your article; for instance, the heading of this subsection will be in a sans serif³ typeface, but that is handled by the document class file. Take care with the use of⁴ the curly braces in typeface changes; they mark the beginning and end of the text that is to be in the different typeface.

You can use whatever symbols, accented characters, or non-English characters you need anywhere in your document; you can find a complete list of what is available in the *\LaTeX User's Guide* [5].

3.1 Macros

If you find you are often using a term that you'd like to typeset in a special way it is a good idea to consider a macro.

² This is the second footnote. It starts a series of three footnotes that add nothing informational, but just give an idea of how footnotes work and look. This is a wordy one, just so you see how a longish one plays out. If a footnote is long you should think hard whether it should be a footnote or a paragraph in the main body of the text.

³ A third footnote, here. Let's make this a rather short one to see how it looks.

⁴ A fourth, and last, footnote.

Command	A Number	Comments
<code>\alignauthor</code>	100	Author alignment
<code>\numberofauthors</code>	200	Author enumeration
<code>\table</code>	300	For tables
<code>\table*</code>	400	For wider tables

Table 1: *Example of a table which spans two columns*

```
% Support Vector Machine
\newcommand{\SVM}{\textsc{svm}}
```

When you use the macro you may need to pay attention to the next character. If the next character is a punctuation you are ok:

```
‘We used an \SVM.’
```

produces: “We used an SVM.”

If the next character is a space:

```
We used an \SVM\ to induce a classifier.
```

The safest way is to put braces around it:

```
We used severals {\SVM}s in cascade.
```

Macros are placed before the `\begin{document}` command.

4. A rather long section w-w-w-w title: An example

Short headings might be a bit too cryptic. Very long headings slow down a reader who is skimming the paper (not too mention that esthetically they don’t look so great).

4.1 Capitalization

The less you capitalize the better. There are some terms that are commonly used by researchers in the field and you will need to mention them. Resist the temptation to capitalize them. The main reasons are again slowing readers and putting emphasis on less important aspects. Terms like ‘information technology’, ‘the web’, etc. are good candidates to keep lowercased. Our recommendation is not to capitalize non-initial content words in section headings. If you have a subheading following a column (:) you may capitalize the next word.

4.2 Indentation

No indentation of:

- the paragraph after section headings (this happens automatically if you use `\section` commands);
- first paragraph on new pages;
- top paragraph of second column;
- paragraph after itemized environments (lists);
- paragraph after figures/tables/equations.

Use `\noindent`.

4.3 Math equations

You may want to display math equations in three distinct styles: inline, numbered or non-numbered display. Each of the three are discussed in the next sections.

4.3.1 Inline (in-text) equations

A formula that appears in the running text is called an inline or in-text formula. It is produced by the `\math` environment, which can be invoked with the usual `\begin{math}... \end{math}` construction or with the short form `$. . . $`. You can use any of the symbols and structures, from α to ω , available in L^AT_EX [5]; this section will simply show a few examples of in-text equations in context. Notice how this equation: $\prod_{i=1}^n p(x_i) = u$, set here in inline math style, looks slightly different when set in display style (see next section).

4.3.2 Display Equations

A numbered display equation—one set off by vertical space from the text and centered horizontally—is produced by the `equation` environment. An unnumbered display equation is produced by the `displaymath` environment.

Again, in either environment, you can use any of the symbols and structures available in L^AT_EX; this section will just give a couple of examples of display equations in context. First, consider the equation, shown as an inline equation above:

$$\prod_{i=1}^n p(x_i) = u \quad (1)$$

Notice how it is formatted somewhat differently in the `equation` environment. Now, we’ll enter an unnumbered equation:

$$\sum_{i=0}^{\infty} x_i + 1$$

Here is a multi-line equation for Naïve Bayes text categorization with unigram features:

$$\begin{aligned} \hat{c} &= \arg \max_{c \in C} P(c|T) \\ &= \arg \max_{c \in C} \frac{P(c) \cdot P(T|c)}{P(T)} \\ &= \arg \max_{c \in C} P(c) \cdot P(T|c) \\ &= \arg \max_{c \in C} P(c) \cdot \prod_{i=1}^n P(w_i|c) \end{aligned}$$

What follows is another numbered equation for the probability of the interpolated bigram language model:

$$p(w_i|w_{i-1}) = \lambda p_{\text{ML}}(w_i|w_{i-1}) + (1 - \lambda) p_{\text{ML}}(w_i) \quad (2)$$

just to demonstrate L^AT_EX’s able handling of numbering.

4.4 Theorem-like constructs

Other common constructs that may occur in your article are the forms for logical constructs like theorems, axioms, corollaries and proofs. There are two forms, one produced by the command `\newtheorem` and the other by the command

`\newdef`; perhaps the clearest and easiest way to distinguish them is to compare the two in the output of this sample document:

This uses the **theorem** environment, created by the `\newtheorem` command:

THEOREM 1. *Let f be continuous on $[a, b]$. If G is an antiderivative for f on $[a, b]$, then*

$$\int_a^b f(t) dt = G(b) - G(a).$$

The other uses the **definition** environment, created by the `\newdef` command:

Definition 1. If z is irrational, then by e^z we mean the unique number which has logarithm z :

$$\log e^z = z$$

There is one other similar construct environment, which is already set up for you; i.e., you must *not* use a `\newdef` command to create it: the **proof** environment. Here is an example of its use:

PROOF. Suppose on the contrary there exists a real number L such that

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = L.$$

Then

$$l = \lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} \left[g(x) \cdot \frac{f(x)}{g(x)} \right] = \lim_{x \rightarrow c} g(x) \cdot \lim_{x \rightarrow c} \frac{f(x)}{g(x)} = 0 \cdot L,$$

which contradicts our assumption that $l \neq 0$. \square

4.5 Citations

Citations to articles [1, 3, 2, 4], conference proceedings [3] or books [6, 5] listed in the References (Bibliography) section of your article will occur throughout the text of your article. You should use BibTeX to automatically produce this bibliography; you simply need to insert one of several citation commands with a key of the item cited in the proper location in the `.tex` file [5]. The key is a short reference you invent to uniquely identify each work; in this sample document, the key is the first author’s surname and a word from the title. This identifying key is included with each item in the `.bib` file for your article. For this article we have provided a sample bibliography in the file `mybiblio.bib`.

You can add your own references to your bibliographical ‘database’ using the notation in the `mybiblio.bib` file. The order in which you specify the references in the database does not affect the order in which they appear in the References section—this is determined by the bibliography stylesheet/template. In the \LaTeX source of your article (your `.tex` file) you need to indicate the stylesheet for the bibliography and the file containing your bibliographical database (here we assume it’s called `mybiblio.bib`). This is achieved by:

```
\bibliographystyle{abbrv}
\bibliography{mybiblio}
```

In the `\bibliography{...}` command you should not specify the `.bib` file extension. The details of the construction of the `.bib` file are beyond the scope of this sample document, more information can be found in the *LaTeX User’s Guide* [5].

This article shows only the plainest form of the citation command, using `\cite{}`; the in-line citation in the previous sentence is achieved by the command `\cite[Lamport:LaTeX]`.

4.6 Tables

Because tables cannot be split across pages, the best placement for them is typically the top of the page nearest their initial cite. To ensure this proper “floating” placement of tables, use the environment **table** to enclose the table’s contents and the table caption. The contents of the table itself must go in the **tabular** environment, to be aligned properly in rows and columns, with the desired horizontal and vertical rules. Again, detailed instructions on **tabular** material is found in the *LaTeX User’s Guide*.

Language	Percentage	Comments
English	51.08%	Lots of folks use English
Chinese	14.73%	Chinese is catching up
Japanese	14.06%	Supplements phone chats
Russian	6.92%	We need Cyrillic fonts

Table 2: Available blog data for ICWSM

Note the sparing use of horizontal and vertical lines.

Referring to figures/tables inline: “In Table 2 we show the popular languages in the weblog dataset.” First, we associate a label with the table:

```
\begin{table}[htb]
\centering
\begin{tabular}{rrl}
...
\end{tabular}
\caption{Available blog data for ICWSM}
\label{tab:weblog-data}
\end{table}
```

Then we use a reference to the table:

“In Table~\ref{tab:weblog-data} we show

The ~ encodes a non-breakable space.

If the figure/table is wide you may consider placing it so that it spans the width of the two columns:

```
\begin{table*}[...] ... \end{table*}
```

As with a single-column table, this wide table will “float” to a location deemed more desirable.

4.7 Figures

Like tables, figures cannot be split across pages; the best placement for them is typically the top or the bottom of the page nearest their initial cite. To ensure this proper “floating” placement of figures, use the environment **figure** to enclose the figure and its caption. This sample document contains examples of `.eps` and `.ps` files to be displayable with \LaTeX .



Fig. 2: A sample black and white graphic (`.eps` format)

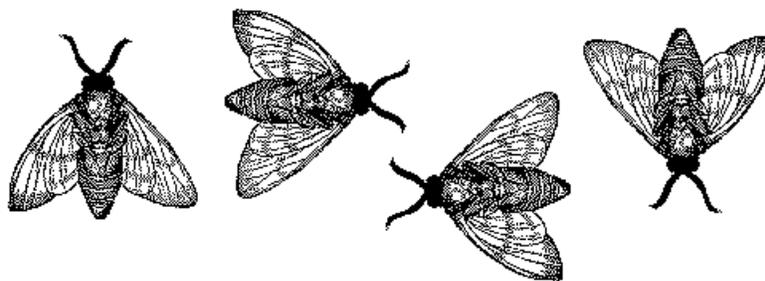


Fig. 1: Black and white graphic (.eps format) spanning two columns

You can associate a label with a figure and refer to the figure using that label as is done for tables.

As was the case with tables, you may want a figure that spans two columns. To do this, and still to ensure proper “floating” placement of tables, instead of enclose the figure and its caption in a `figure` environment use the environment `figure*`:

```
\begin{figure*}[...] ... \end{figure*}
```

Note that either `.ps` or `.eps` formats are used; use the `\epsfig` or `\psfig` commands as appropriate for the different file types.

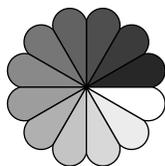


Fig. 3: Black and white graphic (.ps format) that has been resized

In this document Fig.1 has “floated” onto the next page. That is why the first figure that you actually see has a number 2. One way to have the Fig.1 appear physically before the other figures is to place it earlier in the \LaTeX source and let it float onto the desired page. These kind of fine tuning changes should be done when you are finished with the content of the paper.

4.8 Captions

Do not end the captions of tables/figures with periods. One line captions look nice and are easy to read. You may wish not have parentheses italicized in captions by using: `\emph{()}` and `\emph{)}`. Compare the parentheses in the captions of Fig. 2 and Fig. 3.

4.9 Colors

No doubt about it—colors are catchy! RED, BLUE, GREEN, etc. However, you should consider the following:

1. Ultimately the proceedings will be printed in black and white; this will make the colors come out as shades of grey and may not be easily distinguishable from one another.
2. Many reviewers print your submission and read a hard-copy and many printers are set by default to black and

white printing; reviewers may not necessarily browse through the whole submission to decide they should print it in color or on a color printer.

5. Oh, my paper is much longer

After the reviewing process a common request from reviewers is to ask you elaborate on aspects they didn’t quite understand. If you already are at the maximum length of a submission you have a problem. Obviously some bits will have to be omitted. But you have spent so much time on these experiments! And this paragraph/table/picture looks so good! Everybody has the same problem. The organizers have provided a template which is very compact and already allows you to put a lot of information. Think about the main idea you want to convey. Maybe actually you are trying to say more than one thing. Shouldn’t the rest of the important ideas be put in another paper? Maybe for the next ICWSM? Then again, you can have a nice, well structured paper albeit a short one which will interest people and they will contact you for the longer report on which this paper is based. Maybe you can mention the URL for that longer report or cite it.

5.1 Cheating with the length

Here are some obvious ideas to condense your paper, not all of them make the paper look very nice but ...

- using small font for `itemize` environments and maybe some tables too; *like here—see the source*
- put this great but secondary idea in a footnote;
- put a table/figure/equation in one column only;
- redraw figures to take less space;
- control vertical spacing between elements `\vspace*{-.}` or `\[-1.5em]`; *like here—see the source*
- to ensure uniformity of the proceedings please don’t modify the margins or the text height!

A section that has only a single subsection looks very suspicious. You may want to consider using the unnumbered second level heading produced by the command `\subsection*{...}`.

6. Spellcheck the source!

If you are a non-native speaker why don’t you ask a friend of yours who is a native speaker (who went to an exchange in the UK / Ireland / US / Canada / Australia) to have a quick look at your paper. The spellchecker can’t tell you if this phrase sounds rather awkward in English. And it would be nice if the reader flies over your text and not stumble on words and sentences trying to find out what it is you were trying to say.

The ‘basket of words’, ‘forecasting values’, ‘nerve-network’, ‘web-captured’ are usually referred to as the ‘bag of words’, ‘predicting values’, ‘neural networks’, ‘web-crawled’, etc.

7. Page numbering

Do not use page numbering for the camera ready copies. For your submission it is helpful to have page numbers.

Add the command `\pagenumbering{arabic}` above `\begin{document}` if you would like to have the page numbers appear.

8. Available resources

For further examples and detailed descriptions you may wish to consult the ACM *Author’s Guide*: www.acm.org/sigs/pubs/proceed/sig-alternate-v1.htm. Note, however, that the ICWSM style is slightly different and certain commands are not available.

Where can you get this great L^AT_EX typesetting system?

- If you are using UNIX/LINUX chances are L^AT_EX is already installed on your network.
- For PCs here are some (among many) options:
 - cygwin includes L^AT_EX. You need to ensure you have selected it when you installed cygwin. You can always add to an existing installation (www.cygwin.com);
 - MikTeX is another good implementation for Windows (www.miktex.org);
- L^AT_EX is also available for Apple machines (OSX):
 - iTeXMac (<http://itexmac.sourceforge.net>);
 - TeXShop — it won an Apple Design Award in 2002 (<http://darkwing.uoregon.edu/~koch/texshop>);
 - BibDesk — a graphical bibliography manager (<http://bibdesk.sourceforge.net>).

Do bloggers discuss L^AT_EX? Absolutely! Here is one: <http://theappleblog.com/2006/10/05/latex-and-osx>

If you have questions send us an email at info@icwsm.org.

9. Conclusions

This paragraph will end the body of this sample document. Remember that you might still have Acknowledgments or Appendices; brief samples of these follow. There is still the Bibliography to deal with; and we will make a disclaimer about that here: with the exception of the reference to the L^AT_EX book, the citations in this paper are to articles which are used as examples only.

Acknowledgments

This section is optional; it is a location for you to acknowledge grants, funding, editing assistance and what have you. In the present case, for example, the ICWSM organizers would like to thank ACM (Gerald Murray in particular) for allowing us to adapt the ACM stylesheet for our conference.

References

- [1] M. Bowman, S. K. Debray, and L. L. Peterson. Reasoning about naming systems. *ACM Trans. Program. Lang. Syst.*, 15(5):795–825, November 1993.
- [2] J. Braams. Babel, a multilingual style-option system for use with latex’s standard document styles. *TUGboat*, 12(2):291–301, June 1991.
- [3] M. Clark. Post congress tristesse. In *TeX90 Conference Proceedings*, pages 84–89. TeX Users Group, March 1991.
- [4] M. Herlihy. A methodology for implementing highly concurrent data objects. *ACM Trans. Program. Lang. Syst.*, 15(5):745–770, November 1993.
- [5] L. Lamport. *LaTeX User’s Guide and Document Reference Manual*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1986.
- [6] S. Salas and E. Hille. *Calculus: One and Several Variable*. John Wiley and Sons, New York, 1978.

APPENDIX

A. Headings in appendices

The rules about hierarchical headings discussed above for the body of the article are different in the appendices. In the `appendix` environment, the command `\section{...}` is used to indicate the start of each Appendix, with alphabetic order designation (i.e., the first is A, the second B, etc.) and a title (if you include one). So, if you need hierarchical structure *within* an Appendix, start with `\subsection{...}` as the highest level. Here is an outline of the body of this document in Appendix-appropriate form:

A.1 Introduction

A.2 Anatomy of a paper

A.2.1 The authors

A.2.2 The body of the paper

A.3 Type changes and special characters

A.4 Macros

A.5 A rather long section title

A.5.1 Capitalization

A.5.2 Indentationn

A.5.3 Math equations

A.5.4 Theorem-like constructs

A.5.5 Citations

A.5.6 Tables

A.5.7 Figures

A.5.8 Captions

A.5.9 Colors

A.6 Oh, my paper is much longer

A.6.1 Cheating with the length

A.7 Spellcheck the source!

A.8 Page numbering

A.9 Available resources

A.10 Conclusions

A.11 Acknowledgments

A.12 References

Generated by bibtex from your *.bib file. Run `latex`, then `bibtex`, then `latex` twice (to resolve references).