



**Subject:** Publication Information / Workflow

**Contact:** Adam Engelsgjerd  
520.621.2502  
engelsgjerd@u.library.arizona.edu

**Purpose:** This document is intended to document the macro-workflow process for publishing Journal of Insect Science papers. The scope of this document is from the time the paper is sent to the UA Library until it goes live on the Journal's website.

**Date:** This document is current as of November 2002

---

## OVERVIEW

One of the roles the University of Arizona Library plays in publication of the Journal of Insect Science is developing a manuscript from a Rich Text Format (RTF) into a Hyper Text Markup Language (HTML) file and Portable Document Format (PDF) file which are then posted to the website. This document will cover the following areas:

- Personnel
- Equipment
- Software
- Workflow: Overview
- Workflow: RTF to HTML
- Workflow: RTF to PDF
- Working with Figures
- Working with Tables
- Working with Audio / Video Clips
- Working with Large Data Sets
- Submitting to outside agencies, e.g., PubMed

Also included in this CD are:

Appendix A	A sample RTF manuscript
Appendix B	The HTML template used
Appendix C	A completed HTML document
Appendix D	A sample PageMaker document
Appendix E	A sample of a PDF document

## PERSONNEL

Currently, this portion of the University of Arizona Library's involvement in the Journal has one student worker and one full time staff member assigned to it. The average time required to publish a paper is 10 hours (~6 hours of student labor and ~4 hours of the full time employee's labor).

The student employee works on transferring the content from RTF into an HTML template. The full time employee works on transferring content from the RTF into PDF as well as formatting the figures for both HTML and PDF use.

## **EQUIPMENT**

The following equipment is used:

- Windows 2000 workstation  
The full time staff member uses this station to process his portions of the work.
- Linux RedHat 8 workstation  
The full time staff member uses this station to test the site's browser performance.
- Windows NT 4 workstation  
The student employee uses this workstation to process his portions of the work.
- Macintosh OS 9  
The full time staff member and the student use this station to test the site's browser performance

Multiple browser and operating system operability is essential to the success of the Journal. Since it's main readership accesses the content online from all around the world, testing the site in as many different environments as possible a priority. The Macintosh and Linux machines were already available as work began on the Journal. Had they not been available, we may have considered using software packages that test files against different OS / browser combinations. As it stands, this is not necessary.

## **SOFTWARE**

The following software is used:

- Microsoft Word  
This application is used to read the RTF files which include tables and some data sets.
- Microsoft Excel  
This application is used to read some data sets.
- Adobe PhotoShop 5.5  
This application is used to manipulate figures for both the HTML and the PDF production.
- TextPad 4.5  
This application is used as the text-based editor used to work with the HTML of the papers and site in general.
- Opera 5.0  
This application is used in site testing in the Windows and Macintosh environments.
- Adobe PageMaker 7.0  
This application is used to format the PDF version of the paper.
- Adobe Acrobat 5.0  
This application is used to convert the PageMaker document into a PDF file.
- Netscape Navigator 4.x and 6.x  
This application is used in site testing in the Windows and Macintosh environments.
- Mozilla 1.0  
This application is used in site testing in the Windows, Macintosh and Linux environments.
- Internet Explorer 5.x and 6.x  
This application is used in site testing in the Windows and Macintosh environments.

## WORKFLOW: OVERVIEW

In general, the following steps are taken when we publish a JIS article:

- 1 The manuscript, figures, table and data sets are sent via e-mail to the full time UofA Library staff member.
- 2 The manuscript, in RTF, is handed off to the student employee who begins transferring content into the HTML template.
- 3 The full time staff member begins cleaning up the images and re-sizing them to make them ready for publication.
- 4 The full time staff member begins transferring content from the manuscript, in RTF, to PageMaker (which is then transformed into a PDF using Adobe Acrobat).
- 5 The HTML and PDF documents are reviewed for accuracy in both content and display.
- 6 The documents are published (i.e., they are made available online) and the editor, Dr. Hagedorn.
- 7 Corrections are made as necessary either at the request of the authors or Dr. Hagedorn.

The details of this overview process will be discussed in the following sections.

## WORKFLOW: RTF to HTML

When the student employee receives the RTF, his job is to transfer that content into a pre-made HTML template. There are several features of a published paper that he must take into consideration:

- All diacritic marks must be re-coded such that they appear the same in all the browsers listed in the SOFTWARE section.
- All special symbols that the browsers listed in the SOFTWARE section do not support must be replaced with small images, GIFs, of that symbol.
- All textual formatting must be coded in, e.g., italicizing Latin words, applying super or sub script tags to exponents.
- All figure, table or data sets must be prepared to serve on the Internet.
- All references made to the paper's figures, tables or data sets must be linked to the corresponding files.
- All citations made within the body of the paper must be linked to the corresponding reference at the end of the paper.

The content layout and appearance is controlled by the use of Cascading Style Sheets (CSS). The student has access to what is called a Clip Library in TextPad. This allows him to insert relatively long strings of code by simply double-clicking the appropriate clip. For instance, in order to italicize a Latin word in the body of the paper, the following is required:

```
<span class="latin">Drosophila</span>
```

While typing these two pieces of code is not a time intensive process for one word, the nature of these papers is such that Latin appears regularly. Using a clip library in such instance saves us considerable time, and it helps eliminate typing errors in repetitious work.

The figures placed inside figure HTML templates, and the tables are recreated in HTML and transferred over into

page templates.

### **WORKFLOW: RTF to PDF**

Using all Adobe products has made this problematic process significantly easier. There were two major obstacles we faced when first working with JIS content.

The first was formatting the document such that it looked as much like a print publication as possible. We found that Microsoft Word was not a powerful enough application to control the text in the way we wanted it to. Moving over the Adobe PageMaker gave us the control we were looking for.

The second problem was making the figures look good. The HTML version of the paper made it easy for us to produce high resolution and large, in terms of width and height, images. Creating equally high resolution, but smaller in size, width and height, images for the PDF version was substantially more difficult. The primary issue here was the lossy transitions between Adobe PhotoShop, to Microsoft Word and then to Adobe Acrobat. At each step the images lost resolution until in the end, some were unusable. By replacing Microsoft Word with Adobe PageMaker we were able to transfer the images in their PhotoShop format directly into the paper. The only conversion loss experienced, then, is in the conversion of the PageMaker document to the PDF. PageMaker provides a robust Acrobat Distiller (Acrobat) interface allowing us to produce high quality documents.

This process has resulted in large file sizes, but in discussion with Dr. Hagedorn, this is an acceptable trade off for a high quality document.

### **WORKING WITH FIGURES**

Figures are submitted to JIS in a number of different formats. We will accept figures in GIF, JPEG, TIFF, EPS, PNG or PhotoShop format.

One of the main difficulties in this portion of the work is producing a clear image that maintains its integrity in terms of the data represented. Often, submitted figures suffer from pixilation, or fuzziness, around data points, text and images. Whenever possible, the full time staff member will eliminate this pixilation while keeping the data intact. This may involve re-keying in text, e.g., legends, redrawing axis lines and erasing pixilation around images.

The main goal of this work is to get a PhotoShop formatted image that is clear and able to be scaled. Images for the HTML version of the paper are generally re-sized such that their width does not exceed 725pixels. Images for the PDF version are re-sized such that their width does not exceed 3.8inches.

### **WORKING WITH TABLES**

Tables are formatted two different ways, one for the HTML and one for the PDF.

For HTML publication, they are redone using HTML and then placed into a template for display. Doing this allows us to control the look and feel for the table in the online environment.

The PDF version is stripped of its legend comments and saved a separate document. This document is then imported into PageMaker where it is reassociated with the legend information. Doing this allows us to scale the table to fit within a column, or to stretch across the page depending on the table contents' needs.

### **WORKING WITH AUDIO / VIDEO CLIPS**

Multimedia clips are handled in two different ways, depending on the size of the file.

For files that are small enough to be downloaded by the user with the use of streaming technology, we offer them up from Library servers as we would any other file available for download.

For files that are larger, we send them to the Faculty Center for Innovative Instruction, FCII, on the UofA campus. They convert the media into three different streams, RealAudio, Windows Media 1

and Quicktime. They also host the files and we are provided with links which we insert into the HTML and PDF versions of the paper. This aid from the FCII greatly reduces the cost associated with streaming media.

### **WORKING WITH LARGE DATA SETS**

In some instance, authors submit data sets which do not work well in an online environment, e.g., their size does not make them easily usable from a web browser. In such instances, we offer these data sets in their native format as downloads associated with the paper.

### **SUBMITTING TO OUTSIDE AGENCIES**

As the Journal gains recognition and popularity, we have been able to gain acceptance into several scientific resources databases, aggregators and archives, e.g., PubMed.

Such services not only make our content available to a wider audience, but they also provide interlinking between references used in the papers and the full text of the cited information (when available).

Some of these services require the content to be in specific formats in order for them to publish it. In such instances, it is the responsibility of the staff member to modify the content to meet those specifications. Often this involves submitting a textual version of the paper, a PDF version and figures as separate files.