# ExCITE-ing! Build Your Paper's Reference Section Programmatically Using Lex Jansen's Website and SAS®

Louise Hadden, Cormac Corporation

## ABSTRACT

One challenge in writing a SAS® White Paper is creating the perfect reference section, properly acknowledging those who have inspired and paved the way. Luckily, clever use of such tools as Lex Jansen's website, SAS's ability to read in and manipulate varied data sources, and Microsoft Word citation manager, every author can succeed in proper referencing in their white papers. This paper and ePoster will demonstrate how to accomplish this goal.

## INTRODUCTION

The importance of properly citing reference material in a publication cannot be understated. Citation serves a number of purposes: it tells viewers where information can be found and consulted for additional information; it gives appropriate credit to the persons whose words, concepts, and data you are using; and protects you from plagiarism and the risk of having your own research and products invalidated for misusing others' works. PharmaSUG and other groups require that all published work cited in your paper must be listed in a REFERENCES section. This includes both text and visuals (graphics). Special consideration needs to be employed for copyrighted materials in accordance with whether or not you have permission from the copyright or publisher to use the material, either by direct reprint or paraphrasing.

Please note that this paper is specific to SAS White Papers. If you are writing an article for a journal or another organization (APHA, etc.) there will be very specific (and often conflicting from organization to organization) instructions on preparing your works, with regard to how multiple authors are listed, citations, figure and table placements, fonts, and a wealth of other details. The research presented for this paper will inform a tool set for creating reference sections for other types of publications programmatically.

This paper and ePoster are suitable for all skill levels, industries, and job roles. Viewers will gain an understanding of the importance and complexity of proper citations in reference sections as well as resources and tools to help create a robust reference section.

## WHAT DOES A SAS WHITE PAPER REFERENCE SECTION LOOK LIKE?

There are thousands of SAS white papers published, and many of them have excellent reference sections that can be mined both for content and format. However, there are easier ways than Googling, reading, and copying relevant citations by hand.

### SAS USER GROUP MICROSOFT WORD TEMPLATES

You needn't search high and low to find some fantastic instructions for reference sections – very detailed instructions can be found in the Word Template provided in Presenter Resources for this conference. If you haven't looked at the template recently, it is worth a second look as it has been updated for newer versions of Microsoft Word, and for different data sources. It's very tempting to use one of your old papers, or someone else's, as a model, but as the materials being cited evolve, it's worth at least pasting in the reference section of the Word template for up to date models. An important exception is the citation of AI materials, which is covered in this paper. A screenshot of the Reference Section of the Word template for this conference is provided below:

## REFERENCES

Book <Author name: last name, first name (or initials)>. <Publication date>. *<Book title>*. <City, State (abbrev) of publication> : <Publisher name>.

Journal article <Author name: last name, first name>. <Publication date>. "<Article title>." *<Journal title>*, <volume no.;page numbers>.

Article in conference proceedings <Author name: last name, first name>. <Publication date>. "<Article title>." *<Title of proceedings such as Proceedings of the SAS Global 2010 Conference>*, <City, State (abbrev) of publication> : <Publisher name>. Optional: You can add a URL to access available online proceedings. For example: Available at http://support.sas.com/resources/papers/proceedings09/TOC.html.

Website <Author name: last name, first name>. "<Title>." *<Source>*. <Date>. Available at <URL>.

Reference examples:

Book Agresti, A. 2013. *Categorical Data Analysis*. 3rd ed. Hoboken, NJ: John Wiley & Sons.

Journal article Akaike, H. 1979. "A Bayesian Extension of the Minimum AIC Procedure of Autoregressive Model Fitting." *Biometrika*, 66:237–242.

Article in conference proceedings Dorfman, A. H. and R. Valliant. 1993. "Quantile Variance Estimators in Complex Surveys." *Proceedings of the Survey Research Methods Section*, 866–871. Alexandria, VA: American Statistical Association.

Website Federal Reserve Bank of St. Louis. 2012. "Economic Research." Accessed November 7, 2012. http://research.stlouisfed.org.

**Figure 1. PharmaSUG 2025 Word Template**

Items shaded in beige should be removed prior to paper submission, and provide the user with instructions for certain types of citations, with the exception of AI Resources discussed below. It is not comprehensive, but generally provides instructions for most types of documents.

## ADDITIONAL NOTE ON CITING AI RESOURCES

In addition to the References Section from the Word template provided above, if you are using AI resources such as ChatGPT, you can use either APA or MLA style. AI resources are not necessarily retrievable, so should be treated as personal communications which cannot be reproduced. References to AI resources contained within the body of the paper should include the AI Provider (i.e. OpenAI) and the reference date. Full citations in a reference list should also include a description of the prompt and the URL of the AI platform. You can prompt the AI that you are using to provide you with a citation in the appropriate format to make it easy, but provided below are examples of two of the formats that SAS papers customarily accept based on the author's query to Open AI.

*In-text Example :APA guideline*

When prompted with "How many SAS software papers published in 2024 mentioned ChatGPT?" the ChatGPT-generated text provided several detailed citations, primarily at PharmaSUG 2024 and 2025. (Open AI, 2025).

*Reference: APA guideline*

(Open AI, 2025). Response generated by ChatGPT (GPT-4.5-turbo) in response to user query about SAS software papers mentioning ChatGPT. Retrieved from https://chat.openai.com.

*In-text Example: MLA format*

When prompted with "How many SAS software papers published in 2024 mentioned ChatGPT?" the ChatGPT-generated text provided several detailed citations, primarily at PharmaSUG 2024 and 2025. (OpenAI).

*Reference: MLA format*

OpenAI. *ChatGPT* (GPT-4.5-turbo). Response to a query about SAS Software papers mentioning ChatGPT. 18 May 2025. *ChatGPT*, https://chat.openai.com/.

## ADDITIONAL RESOURCES FOR HOW-TOS ON CITATIONS

SAS User Group templates and advice on citing AI resources used are not inclusive of other types of writing and may be out of date due to ongoing changes in technology and what needs to be cited – the concept of digital copyrights is fairly fluid at this point. Listed below are several other resources that may be helpful in writing your SAS white paper.

### SAS Editorial Guidelines

SAS publishes Editorial Guidelines which are updated on a regular basis. These guidelines mostly cover how SAS's copyrighted information should be treated, but do not comprise a full set of citation rules. Note that you ALWAYS need to consult the guidelines your university, your employer, or potential publisher have set forth.

https://www.sas.com/content/dam/SAS/documents/legal-employee-services/legal/editorial-guidelines.pdf

### Commonly Used Instruction Sets for Citations

Citation guidelines in SAS white papers include both in-text citations (brief parenthetical mentions of the resource) within the paper and a stylized, more complete references section at the end, as noted above in the section on AI citations. Using these guidelines ensures that all resources are properly credited and viewers can find the details on cited information. Most commonly used instruction sets are APA (American Psychological Association), MLA (Modern Language Association), and Chicago/Turabian (used in business, history, fine arts, etc. – not necessarily relevant for SAS White Papers.) The SAS User Group Word template states that SAS authors may use the APA or MLA systems.

### Microsoft Word Citation Management

The process of writing SAS papers has evolved in strange and wonderful ways since my first SAS paper in two column format and just a few indications of the structured Word templates we have now. We have the editor, we have

### ANTI-Plagiarism Tools

It may seem like all the citation rules are overkill, but having the framework in place does establish your credibility and reliability of your work. Another aspect of this is avoiding plagiarism – which is actually hard to do with SAS white papers, where we all learn from one another. There are number of anti-plagiarism tools available in both software and website form. These tools and sites work by comparing the text in your work with texts in large online databases. Some popular options include: Scribbr, Grammarly, Copyleaks, Quillbot, iThenticate, DupliCheckr, and PlagScan. At the very least, Microsoft Word Users can access the built-in editor to "Check for similarity to online sources."

In addition to protecting yourself from inadvertently using others' work, you can protect your own papers by securing your PDFs with Adobe, and/or opening your word document and editing your properties (metadata) of your Word or PDF documents. The important thing with metadata is to add keywords in the Tag section so that your paper can be recognized when people electronically search for a given topic or element in papers, on Lex Jansen's web site described below, or elsewhere. Search engines read the tags.
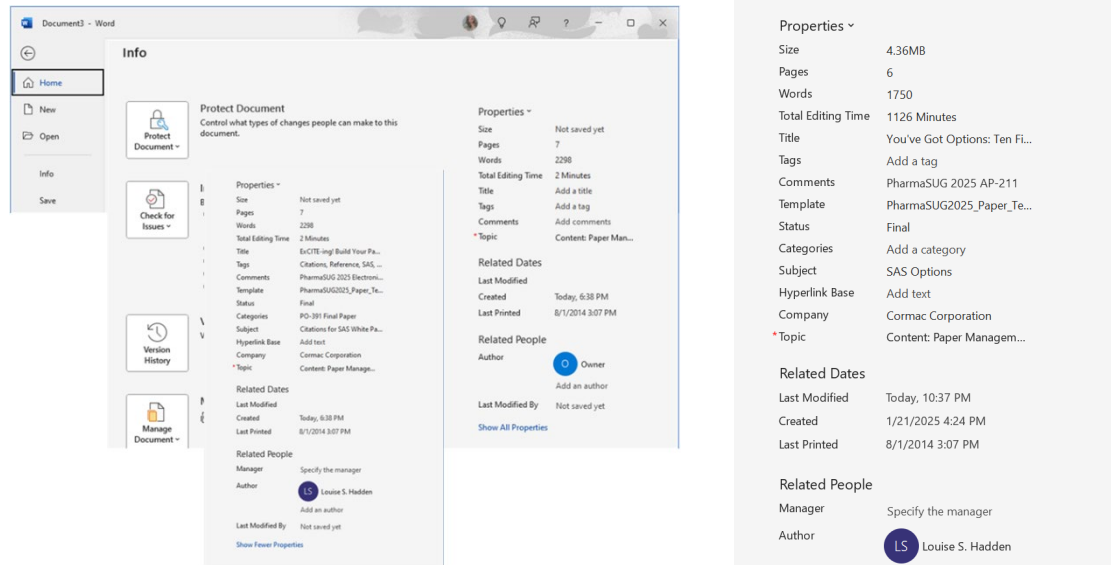
**Figure 2. Entering Paper Metadata in MS Word and PDF**

## LEXJANSEN.COM

Lex Jansen's web page LexJansen.com is my primary go-to with any SAS question or problem I may have – there are many other resources, Google, listservs, SAS communities, friends – but Lex's site is particularly convenient, with a number of different pages with different functionalities. Lex is a former SAS employee, and many years ago undertook the collection of proceedings from SAS User Group and other related conferences, including scanning by hand for earlier years. It's a mammoth undertaking and is constantly being nurtured with the yearly influx of new SAS white papers, some, including this paper, describing this history in the making.

Early in my SAS career, I was introduced to Lex and his efforts, and his website and sascommunity.org, a wiki-based website, and they became my primary sources of SAS information. A decade ago, I asked Lex if he would be willing to add XML as an output destination to save searches, et voila, XML happened and JSON was not far behind. This paper and e-poster will discuss using saved search results and creating a reference section for SAS white papers electronically, but Lex's site can entertain a number of other use cases, such as topic frequency analyses  (Hughes, Proceedings of WUSS 2023).

## MAP OF PAGES ON LEXJANSEN.COM



**Figure 3. Word Cloud on Lex Jansen's Web Page**

## Top Level

The main page of Lex Jansen's conference proceedings site provides you with a Google enhanced search field looking at over 38,840 SAS white papers and counting, user group level selections, selected author bibliographies and more! https://www.lexjansen.com/
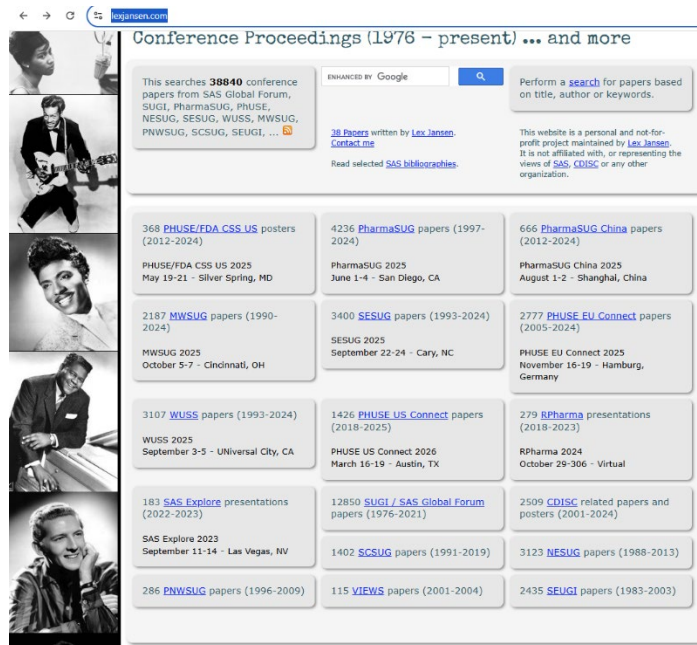


**Figure 4. Lex Jansen's Top Level Page**

## SAS Search

At the top right of the main page, clicking on the "perform a search" link will bring you to the SAS search page. Here, you can see individual user group pages (also seen on the main page), and you can search by title/keywords and/or author(s).
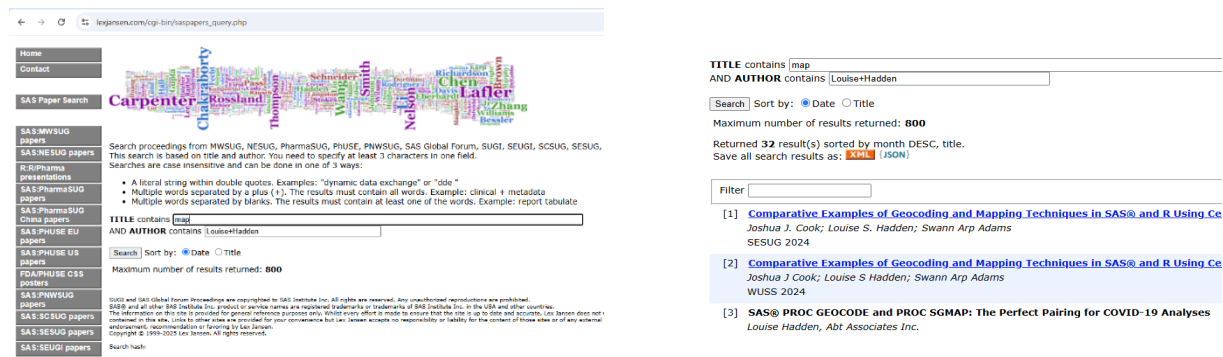


**Figure 5. Lex Jansen's SAS Search Page**

The results of a search are shown above right – I have 32 papers with the keyword maps in the title. Note that you can save the search results as XML or JSON with a simple click. You can also click on any of the papers listed to download them individually, or right click and save the link for later use. XML and JSON downloads will be explored later. https://www.lexjansen.com/cgi-bin/saspapers_query.php

## USER GROUP LEVEL (I.E. PHARMASUG)

You can reach this level from the main page or the SAS search page by clicking on one of the User Group selections, in this case PharmaSUG. https://www.lexjansen.com/pharmasug/



**Conference Proceedings: Pharmaceutical Software Users Group 2024**

**May 19-22, 2024, Baltimore, Maryland, 206 papers**

PharmaSUG 2024  PharmaSUG 2023  PharmaSUG 2022  PharmaSUG 2021  PharmaSUG 2020  PharmaSUG 2019  PharmaSUG 2018  PharmaSUG 2017  PharmaSUG 2016  PharmaSUG 2015  PharmaSUG 2014  PharmaSUG 2013  PharmaSUG 2012  PharmaSUG 2011  PharmaSUG 2010  PharmaSUG 2009  PharmaSUG 2008  PharmaSUG 2007  PharmaSUG 2006  PharmaSUG 2005  PharmaSUG 2004  PharmaSUG 2003  PharmaSUG 2002  PharmaSUG 2001  PharmaSUG 2000  PharmaSUG 1999  PharmaSUG 1998  PharmaSUG 1997

[Original source: https://pharmasug.org/conferences/pharmasug-2024-us/conference-proceedings/]

**Pharmaceutical Software Users Group 2024: Best Papers and honorable mentions**

Build Your Own PDF Generator: A Practical Demonstration of Free and Open-Source Tools
  James Austrow, Cleveland Clinic (*Best Paper* in Advanced Programming)
Replicating SAS® Procedures in R with the PROCS Package
  David Bosak, r-sassy.org (*Best Paper* in Advanced Programming)
A quick guide to SDTM and ADaM mapping of liquid Oncology Endpoints.
  Swaroop Kumar Koduri, Ephicacy Lifescience Analytics Pvt Ltd; Shashikant Kumar, Ephicacy Lifescience Analytics; Sathaiah Sanga, Ephicacy Lifescience Analytics (*Best Paper* in Data Standards)
Quarto 1.4: Revolutionizing Open-source Dashboarding Capabilities
  Joshua Cook, University of West Florida (UWF); Kirk Paul Lafler, sasNerd (*Best Paper* in Data Visualization and Reporting)
Leadership Lessons from Another Life: How my Previous Career Helped Me as a Statistician
  Christiana Hawn, Catalyst Flex; Lily Ray, Catalyst Clinical Research (*Best Paper* in Leadership Skills)
A Practical Approach to Automating SDTM Using a Metadata-Driven Method That Leverages CRF Specifications and SDTM Standards
  Xiangchen Cui, Crisprtx Therapeutics; Min Chen, CRISPR Therapeutics; Jessie Wang, CRISPR Therapeutics (*Best Paper* in Metadata Management)
Win a PS5! How to Run and Compare Propensity Score Matching Performance Across Multiple Algorithms in Five Minutes or Less
  Catherine Briggs, SAS; Sherrine Eid, SAS Institute; Samiul Haque, SAS Institute; Robert Collins, SAS Institute (*Best Paper* in Real World Evidence and Big Data)
Integration of SAS GRID environment and SF-36 Health Survey scoring API with SAS Packages
  Bart Jablonski, yabwon (*Best Paper* in Solution Development)
A Tool for Automated Comparison of Core Variables Across ADaM Specifications Files
  Amy Zhang, Merck & Co.; Huei-Ling Chen, Merck & Co. (*Best Paper* in Solution Development)
Dealing with Missing Data: Practical Implementation in SAS and R
  Isabella Wang, Eli Lilly & Co.; Jin Xie, Eli Lilly & Co.; Lauren George, Eli Lilly & Co. (*Best Paper* in Statistics and Analytics)
A fear of missing out and a fear of messing up : A Strategic Roadmap for ChatGPT Integration at Company Level
  Kevin Lee, Karuna Therapeutics (*Best Paper* in Strategic Implementation & Innovation)
A Programmer's Insight into an Alternative to TQT Study Data Submission
  Hiba Najeeb, Vertex Pharmaceuticals; Raghavender Ranga, Vertex Pharmaceuticals (*Best Paper* in Submission Standards)
Simplifying Edit Check Configuration
  Vandita Tripathi, Ms; Manas Saha, TCS (*Best Poster* in e-Posters)

**Figure 6. Lex Jansen User Group Level**

## MISCELLANEOUS LEVEL

On this level, you can find a hidden treasure: special SAS biographies created by the Macro Maven, Ron Fehd. This level is found by clicking on SAS Biographies on the main level. You can also find a selection of Lex Jansen's papers and a link to Troy Martin Hughes' paper.

https://www.lexjansen.com/sas-bibliographies/

https://www.lexjansen.com/sas-bibliographies/Hughes-2024_wuss-216_48-year-analysis.pdf

https://www.lexjansen.com/papers/

It is possible to extract an amazing amount of data, papers, etc. in an instant from Lex Jansen's site using the tools provided, and via web scraping. Troy Martin Hughes uses Python and Beautiful Soup to scrape the site and perform a number of analyses on SAS user group papers / presentations, and Bart Jablonski has Packages that automate the tasks using SAS. Please see the reference section for more information on their related work.

## ADDITIONAL SOURCES OF BIBLIOGRAPHICAL MATERIALS

There are myriad sources for literature searches and citations on the internet, including other search engines such as Google, sites such as Github, SAS Communities, Stackoverflow, etc. Most allow the copying or scraping of materials to create a reference section.

## REFERENCE REPORTING DATA BASE CONSTRUCTION

Once electronic data collection of reference materials has been accomplished, it is time to identify the elements of the citation of SAS White Papers. Reviewing the citation formats discussed above will provide the details of the separate elements that need to be melded into a reference, and the information from Lex Jansen's site (multiple tables!) will be explored and mapped. We have explored the formats for citing various types of articles, books, etc. above, but for the purposes of demonstration, we can use existing citation models from the PharmaSUG 2025 Paper Template.

## ELEMENTS OF THE CITATION

"Article in conference proceedings <Author name: last name, first name>. <Publication date>. "<Article title>." *<Title of proceedings such as Proceedings of the SAS Global 2010 Conference>*, <City, State (abbrev) of publication> : <Publisher name>. Optional: You can add a URL to access available online proceedings. For example: Available at http://support.sas.com/resources/papers/proceedings09/TOC.html."

From the snippet above, we can determine that we need the following elements:

- Author name, formatted as Last, First, followed by a period.

- Publication Date, formatted as Year enclosed in parentheses, followed by a period.

- Title, enclosed in double quotes, followed by a period. Optional italics.

- Proceedings title or Conference name, formatted as "Proceedings of" conference name and followed by "Conference" and a comma.

- City and State of the Conference, followed by a period.

- (Optional) "Retrieved from " link to paper

## AVAILABLE DATA

There are three basic methods of obtaining data from Lex Jansen's site: web scraping, a JSON download, and an XML download. All three result in multiple tables, and all three require post-processing to get the necessary elements of a reference section. We will focus on the JSON method for the purposes of this paper, noting that all three methods are basically similar.

A search is performed on the SAS search page of Lex Jansen's site, searching by author for Bart Jablonski. A JSON file is downloaded. The following program is run:
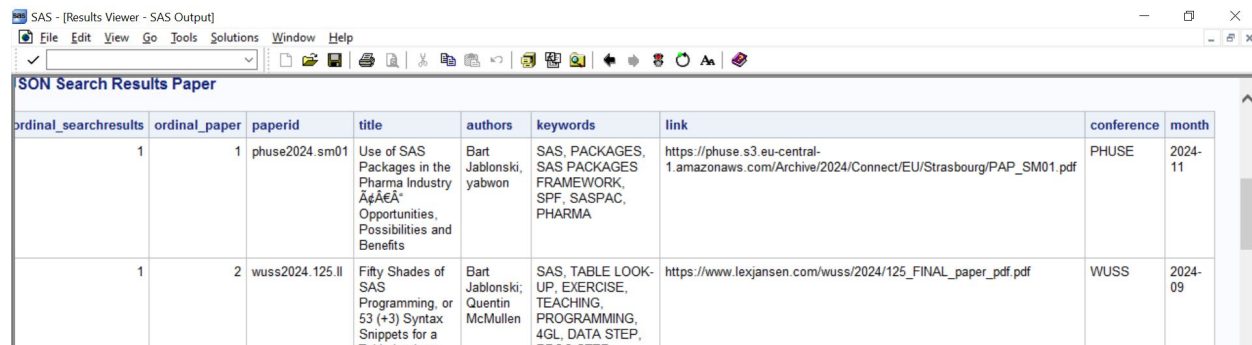
```
%let progpath=C:\Papers\391_ExCITE-ing!;

libname dd "&progpath";
filename jsonfile "&progpath.\searchresults.json";
libname jsonfile json fileref=jsonfile noalldata;
run;


proc copy in=jsonfile out=dd;
run;
```

**Program 1. Code Snippet for JSON read-in**

The process results in three separate tables: a paper download file, a search result file (this contains the search date, which you can store for reference), and a search result paper file. The search result paper file contains the bulk of the information we need.



**Figure 7. Search Results Table Snippet**

Conference and Month comprise part of "Proceedings of" and "Publication Date" ; and Title gives us Title; Authors gives us Authors; and Link will give us link. We still need location information for the conferences

and we can obtain this information from Lex's site as well, from an XML file available on the main page of Lex's site. https://www.lexjansen.com/sasproceedings.xml
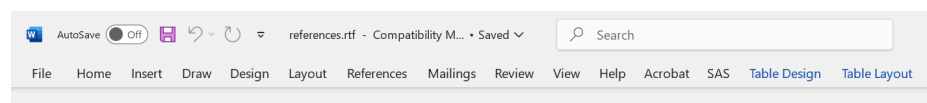


**Figure 8. Locations Snippet**

We now have all the elements needed. Locations can be added to the citation information via merging on conf_url in the location table and scanning the first chunk of paperid in the search results table.

## OUTPUT A REFERENCE SECTION USING SAS

Features of the SAS program to create a reference section include:

- Merge by conf_url (created in the search results table)
- Precleaning (dirty Unicode data)
- Preprocess all segments of the citation, particularly name, using SAS functions
- Create a very long string which includes all elements
- Sort output by desired keys (generally last name of the first author, first name of first author, title)
- Output to RTF for an easy drop into the reference section

Full code, the paper, and presentation will be available in my GitHub page following the conference.



*Prepare a Reference Section for a SAS Paper*

Jablonski, Bart. (2023). "A SAS Code Hidden in Plain Sight". Proceedings of Western Users of SAS Software, San Diego, California. Retrieved from https://www.lexjansen.com/wuss/2023/WUSS-2023-Paper-189.pdf

Jablonski, Bart. (2023). "Share your code with SAS Packages - a Hands-on-Workshop". Proceedings of Western Users of SAS Software, San Diego, California. Retrieved from https://www.lexjansen.com/wuss/2023/WUSS-2023-Paper-208.pdf

Jablonski, Bart. (2024). "A Gentle Introduction to SAS Packages". Proceedings of Pharmaceutical SAS Users Group, Baltimore, Maryland. Retrieved from https://www.lexjansen.com/pharmasug/2024/HT/PharmaSUG-2024-HT-111.pdf

**Figure 9. References RTF Output File Snippet**

## CONCLUSION

Lex Jansen's web site has been a wonderful resource for me for literature searches, trouble-shooting, and writing my own reference sections for SAS white papers. I hope I inspire others to take advantage of the inspiration and convenience.

## REFERENCES

Hughes, Troy Martin. (2023). "Who's Bringing That Big Data Energy?  A 47-Year Longitudinal Analysis of 30,000 Presentations in the SAS® User Community To Elucidate Top Contributors and Rising Stars". *Proceedings of the Western Users of SAS Software 2023 Conference,* San Diego, California. Retrieved from https://www.lexjansen.com/wuss/2023/WUSS-2023-Paper-204.pdf

Hughes, Troy Martin. (2024)." Who's Still Bringing That Big Data Energy?  Leading a SAS® Conference by Leveraging a 48-Year Longitudinal Analysis of 30,498 Presentations in the SAS User Community". *Proceedings of the Western Users of SAS Software 2024 Conference*, Sacramento, California. Retrieved from https://www.lexjansen.com/sas-bibliographies/Hughes-2024_wuss_216_48-year-analysis.pdf

Jablonski, Bart. (2023). "Share your code with SAS Packages - a Hands-on-Workshop". *Proceedings of Western Users of SAS Software 2023 Conference,* San Diego, California. Retrieved from https://www.lexjansen.com/wuss/2023/WUSS-2023-Paper-208.pdf

## ACKNOWLEDGMENTS

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Louise S. Hadden
louisesquibbhadden@gmail.com

Any brand and product names are trademarks of their respective companies.