

5-31-2012

Making Bibliographic Researchers More Efficient: Tools for Organizing and Downloading PDFs, Part 2: QUOSA

Mary Moore PhD

University of Miami Miller School of Medicine, mmoore@med.miami.edu

Kimberly A. Loper

University of Miami Miller School of Medicine, kloper@med.miami.edu

David Goolabsingh

University of Miami Miller School of Medicine, dgoolabsingh@med.miami.edu

JoAnn Van Schaik

University of Miami Miller School of Medicine, Jvanschaik@med.miami.edu

Recommended Citation

Moore, M.; Loper, K.A.; Goolabsingh, D.; and Van Schaik, J. Making Bibliographic Researchers More Efficient: Tools for Organizing and Downloading PDFs, Part 2: QUOSA. 9(2):114-125, 2012. DOI:10.1080/15424065.2012.680314

This Article is brought to you for free and open access by the Department of Health Informatics at Scholarly Repository. It has been accepted for inclusion in Faculty Research, Publications, and Presentations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.

eJOURNALS FORUM

Mary Moore and Suzetta Burrows, Column Editors

*Making Bibliographic Researchers More Efficient: Tools for Organizing and
Downloading PDFs, Part 2: QUOSA*

Mary Moore

Kimberly A. Loper

David Goolabsingh

JoAnn Van Schaik

ABSTRACT. The development of integrated software programs should help assist researcher-writers in completing more routine duties, such as bibliographic editing. In addition, some programs also help researcher-writers with the analysis and synthesis of search results. Such programs can save time and result in new discoveries. Part 1 of this exploration, published in the last issue, covered iCyte[®], Mendeley[™], Papers, PDF Stacks, PubGet PaperPlane[™], WizFolio, and Zotero[™].¹ Part 2, published here, covers a robust and highly integrated product, QUOSA[™], including an in-depth product performance evaluation. Questions considered in this column are: What types of programs exist to enhance the productivity of researcher-writers? What are their purposes? How do these programs operate? Do they always behave as promised?

KEYWORDS. Bibliographic research, citation analysis, QUOSA, research, scholarly communication

Authors. Mary Moore, PhD, (mmoore@med.miami.edu) is the corresponding author and is Chair, Department of Health Informatics, and Executive Director, Medical Libraries and Biomedical Communications at the University of Miami Miller School of Medicine, P.O. Box 016950 (R-950), Miami, FL 33101.

Kimberly A. Loper, MLIS, (kloper@med.miami.edu) is Assistant Vice Chair, Department of Health Informatics, Administration, and Special Projects and Digital Initiatives Librarian at the University of Miami Miller School of Medicine, P.O. Box 016950 (R-950), Miami, FL 33101.

David B. Goolabsingh, MLIS, (dgoolabsingh@med.miami.edu) is Director for Information Technologies, Department of Health Informatics at the University of Miami Miller School of Medicine, P.O. Box 016950 (R-950), Miami, FL 33101.

JoAnn Van Schaik, MLS, (jvanschaik@med.miami.edu) is Vice Chair, Department of Health Informatics and Director, Medical Libraries Operations, at the University of Miami Miller School of Medicine, P.O. Box 016950 (R-950), Miami, FL 33101.

Author credits: Mary Moore, PhD, idea and development of the overall plan and organization of the paper, and authored following sections: Abstract, Introduction, How QUOSA has been used, and Summary; Kimberly A. Loper, MLIS, authored the in-depth description of QUOSA and formatted the paper, and constructed the bibliography; David Goolabsingh, MLIS and JoAnn Van Schaik, MLS., created the performance review of QUOSA searches and analyzed the results.

Acknowledgements: Yanira (Jenny) Garcia-Barcena, MLS, Calder Reference Librarian, created the search strategies for the performance review. Eszter Hars, PhD, Customer Development Scientist with QUOSA, provided prompt responses to questions.

MAKING BIBLIOGRAPHIC RESEARCH MORE EFFICIENT

There is substantial opportunity for increased efficiency when conducting bibliographic research and writing scholarly papers. The process for conducting bibliographic research in the days before e-journals was described in the last issue in Part 1¹ of this column.

Today search engines are user friendly and many journal articles are available online, yet there still is room for improved efficiency. An ideal system would allow a researcher-writer to click just one button after performing a bibliographic search and download all selected articles at once. From there, the article PDFs could be placed into an electronic folder on the researcher's computer desktop, and the full text content searched, sorted and categorized electronically. Researchers could enter one search string to find all articles with relevant terms, which would be highlighted in each article. Relevant passages could be color coded electronically, and when a researcher wanted to see all the passages on a particular topic from all the articles, they could be quickly reviewed and the researcher could add electronic annotations next to the relevant passages in the article. From that point, relevant passages could be cut and pasted into the new article, automatically importing the citations from where the quotations originated. Researchers should be able to initiate a text analysis to identify the most commonly recurring phrases in a group of articles, helping to identify common themes or emerging trends in the literature. As the author's article took form, the author, with the help of the integrated program, should be able to

identify the name of the journal to which the article would be submitted, considering elements such as content, types of articles solicited, journal impact factor, and open access/public access factors. Another aspect of the integrated program would check the article against the journal's style guide, including citation formatting. In a final step, an integrated program would then check each line of the submitted article against the content of the articles and the Web for accidentally uncited quotations. The author's productivity would be increased, opportunities for creative analysis and synthesis would be increased, and unintentional plagiarism would be avoided.

Unfortunately, the authors of this column were unable to find any existing products with all of this functionality. Products previously described¹ provided limited aspects of some features, such as the ability to download, organize and annotate a PDF. One product, however, comes closer than the others to reaching the ideal.

TOOLS TO ASSIST IN RESEARCHER-WRITER PRODUCTIVITY: QUOSA

QUOSA[®] Information Manager (QUOSA, Inc., Waltham, Massachusetts) is more functional than the other products reviewed in Part 1.¹ *QUOSA* stands for *QU*ery, *O*rganize, *S*hare and *A*nalyze (four of the program's five key capabilities – with *S*ave being the fifth) and is designed to aid in text-searching. In addition to being able to take a search and download multiple PDFs with one click, QUOSA allows the researcher to organize, annotate and search within multiple PDFs. It can also be used for cluster analysis of commonly occurring terms in all PDFs from a search. Development of QUOSA began in 1992 and the first version was launched in 2002.² The product is currently available institution-wide at over 300 organizations in more

than 35 countries.

Once a downloadable plug-in is installed, the product uses Internet browsers to interface with search engines from PubMed™, Scopus™, NIH RePORTER™, and Google™, among others (called "channels" by QUOSA). With the click of one button, QUOSA downloads and organizes up to 500 search results as full-text PDFs, HTML files, or as abstracts only, as allowed by the institutional library's subscriptions. These downloaded results are stored on the user's computer hard drive so that the files can be accessed at any time.

QUOSA is launched from an icon on the desktop and consists of a *menu bar* and three panes: the *Browser Pane*, *Results Pane* and *Organizer Pane*. QUOSA has two different viewing options. The *Express View* is the default view when QUOSA opens. In this view, the toolbar icons are labeled. Selecting *Advanced View* will remove the toolbar icon labels on both the *Browser Pane* and *Results Pane*, leaving just the correlating icons. *Advanced View* adds *Concepts4Clustering* and *Terms4Clustering* tabs to the *Organizer Pane* and adds the URL of the current results to the *Browser Pane Toolbar*.

[PLACE FIGURE 1 HERE]

Legend Figure 1: QUOSA Interface divided into menu bar 1, Browser Pane 2, Results Pane 3, Organizer Pane 4 and the View options selection area 5.

Menu Bar

The *Menu bar* is anchored along the upper region of the screen and is comprised of nine drop down menus. The first three: "file," "edit," and "view" allow for many common functions

such as folder management, backup tools, file sharing, printing, quit and close.⁴ *Commands* is the next drop down menu on the menu bar. It contains selections that allow for *searching*, *selecting*, and *exporting* results. These are options that are also available on the *Results Pane* menu bar. One of the more beneficial selections available under the *Commands* drop down menu is the *Sync & Link* selected articles function. This function allows QUOSA to export all the user's downloaded results with just one click. This feature works with several citation management software products including EndNote™, Reference Manager™, ProCite™, and RefWorks™.³

The *Settings* drop down menu allows the user to customize the appearance of the *Results* in the *Results Pane*. Viewed and unviewed *Results* can be differentiated by using a variety of font sizes, colors and styles. Here the user can also highlight relevant text sections which help to quickly locate the user's search terms within each result. An available add-on feature in the *Settings* menu is *Rightsphere*. When *Rightsphere* is active, a *Rights* column that integrates with the Copy Clearance Center will be displayed in the *Results Pane* and, with the click of a button, *Rights* allows the user to check copyright status of each downloaded article.³ *Channels* is where database selection occurs. Basic channels/database options are OvidSP, PubMed, Google, Google Scholar, and RePorter. Premium channels (Web-of-Knowledge, Embase and Scopus) can be added depending on institution subscriptions. Here the user can also select a federated search. This allows the system to search multiple "basic" databases with one search.

Other available options on the *Menu bar* include the *Help* menu, the *Advanced View/Express View* button, *Results Button*, *Browser button* and *Both button*. *Help* contains a searchable online manual with information on how to use QUOSA and is where to go to contact QUOSA support. The *Advanced View/Express View* button toggles between the two different

views. The *Results* button displays the *Results Pane* only. The *Browser* button displays the *Browser Pane* only. The *Both* button displays both the *Browser Pane* and the *Results Pane* simultaneously.

Browser Pane

One of the benefits of QUOSA is that it does not require learning new search strategies because it has a built-in Internet Browser. The *Browser Pane* is located on the bottom portion of QUOSA and has a *toolbar* at the top of the pane. In the *Express View*, the *toolbar* has labeled icons. In the *Advanced View* the labels are removed from the icons, but an address field is added to the toolbar. Once a *Channel* is selected on the *Menu Bar*, the *Browser Pane* is filled with that Website, which is searched using the same strategies used if entered into any Internet browser.

When the search is complete, the results will be shown. In *Browser Pane* the user reviews returned documents and selects those that should be retrieved by QUOSA. On the *Browser Pane toolbar*, the user would type in the desired number of results to retrieve (up to a maximum of 500), and use the drop down arrow to select the preferred file type. The options are PDF, HTML or abstract only. If the format the user chooses is not available, QUOSA will automatically download the next “best” available type. If the user clicks the Σ (Retrieve) button on the *toolbar*, QUOSA will download the selected articles to the *Results Pane*.

Another item on the toolbar is the *Citations* button. Using the *Citations* tool allows users to simultaneously retrieve and store selected items in the *Results Pane* and export citations to the user’s citation manager. In addition to the original citation information, QUOSA creates two links for every result during export. One points back to the article stored in QUOSA and the other points back to the source of the article.⁵

Results Pane

The *Results Pane* is above the *Browser Pane*. After selecting articles and clicking the Σ (Retrieve) button, articles are displayed in the *Results Pane*. Like the *Browser Pane*, the *Results Pane* also has a tool bar to aid navigation. Below the *toolbar* are column headers that give information on retrieved articles. This includes *Authors*, *Title*, *Type of file*, *Source*, *Date published*, *Search Engine* retrieved from, *file Size*, *Search Engine* rank and date *Downloaded*. Users can sort the results based on *Header* by clicking (and clicking again to sort in the opposite order). *Headers* can be reordered by clicking on the one to select and then while selected move by dragging it to the desired locations.

The *Tool bar* allows users to move forward or back through the results by clicking on the *Open Next* or *Open Previous* icons. *Save* is used to save a result in a folder on the *Organizer Pane*. Also on the *Tool bar* is the *Search in Results* icon. *Search in Results* allows the user to narrow results by entering terms, displaying only those results in which the term is found. This search can be done as a quick search or as an Advanced search. Quick searches search all fields in the results set, including the full text of documents. Advanced searches offer searchers the ability to choose the field where the term is to be searched, along with the ability to search all files and folders at the same time rather than just the displayed set of results. For quick review, the *Highlight* button allows users to choose different colors for searched terms. These will be highlighted in the display screen.

Organizer Pane

The *Organizer Pane* is sorted into tabs. In the *Express View*, *My Article Organizer* and *Document Summary* tabs are available. In the *Advanced View*, *Concepts4Clustering* and *Term4Clustering* tabs are activated. “The Organizer Pane shows your results saved in a personalized folder structure.”³

The default view for the *Organizer Pane* is to have the *My Article Organizer* tab open. The folders available here are *Inbox*, *My Searches*, *My Alerts*, *My Folders*, *My Citations*, and if activated, *Virtual Libraries*. All imported items will be placed in the *Inbox*. Every day QUOSA is used, a folder with the date will be created under the *My Searches* folder. All searches conducted on a particular day will be listed in that folder. Once a search has been created, QUOSA can automatically rerun the search using the *Alerts* function. If the user sets a search to run an alert, QUOSA will deliver only the new results posted to the database after the prior search.⁹ These results can be found in the *My Alerts* folder. *My Folders* is used to organize the user's results and the *My Citations* folder helps organize those results that have linked to the user's citation manager. The *Document Summary* tab displays all relevant searched terms. The user can move quickly through the document to each relevant term by clicking the appropriate line in the *Document Summary* tab.

Concepts4Clustering is a tab that is available only in the Advanced View. When the user opens the *Concepts4Clustering* tab, an automatic analysis of all search results is created. After this analysis, results are grouped by repeated terms or concepts and are displayed in the tab. This might give searchers a better understanding of each result's content, and allows users to quickly browse results that could be more similar.

QUOSA APPLICATIONS

There are many examples of how QUOSA might be used, in addition to the most obvious ways.

For example:

- After downloading articles from PubMed or Scopus, the “Concepts for Clustering” feature might be used to help make sense of a topic or field of study, when a researcher is new to the topic. Seeing the common terms used across a number of articles can help a

researcher-writer understand common sub-topics addressed in a topic or field of study.

This feature can also help a writer determine if certain topics were left unaddressed in the article the writer is preparing.

- The optional installation of QUOSA onto a user's desktop allows users to download numerous articles onto their computers, allowing users to access the articles anywhere. This option has multiple applications, for example, reading articles on an airplane or on a cruise ship, or anywhere network connections might be limited or unavailable. It also allows the user to compile personal computer collections of often cited and used articles.
- In times of disaster or emergency, a savvy physician who has downloaded PDFs with current clinical protocols or guidelines would not have to be as concerned about lack of live network access. This might also be a good solution for healthcare providers who travel to practice medicine in isolated or rural areas that might have limited access to online resources.
- Some users have reported using QUOSA on their laptop computers before they attended conferences. They prepared for the conference by downloading articles by the conference speakers and selected which sessions to attend by scanning the articles the speakers had written before the lectures. Reportedly, being able to access the articles quickly has enhanced the conference experience.
- Experienced writers have also reported using the feature that links QUOSA with a reference manager to take older bibliographies from works written early in the career of the writer and then download electronic copies of articles that were previously cited. Since many researchers work in the same field over their careers and therefore often cite

the same articles repeatedly, this can expedite writing the Introduction and Background section of studies that have evolved from previous studies.

- In some cases, researchers have used QUOSA specifically as part of the research methods in a study. For example, three of the authors of this column conducted a study of how people might search Google for self-diagnosis, using QUOSA to download and analyze relevant Websites.⁶ There are a few examples from a Scopus search on how QUOSA has been used for research. Beyrer et al⁷ describe the use of QUOSA in conducting a systematic review. Xu et al also describe the use of QUOSA to “enable automated full-text searches of any number of drugs AND any number of classical terminologies for liver injury ...”⁸ Chaussabel describes how software can be used for biomedical literature mining and semantic analysis.⁹ However, if researchers are to depend on QUOSA for systematic analysis, an understanding of how it practically performs is necessary.

PERFORMANCE REVIEW OF QUOSA

Study Introduction and Research Question

Does QUOSA perform as promised? Researchers at the University of Miami had previously reported that they wondered whether all search results were successfully exported from PubMed into QUOSA, whether all full text articles that were available from university subscriptions appeared in QUOSA results, and whether the "search within" function of QUOSA was accurate. Researchers also described how, on occasion, they had typed a term that had resulted in completely unrelated search results in QUOSA.

It should be noted that researchers had indicated that in every case where they had been able to document the problem and relay it to QUOSA, the problem had been resolved. In some

cases problems were caused because researchers had not installed a product update. This study took a systematic approach to determining whether QUOSA worked as predicted.

Methods

The subjects of the five QUOSA searches conducted for this study reflect the strengths and areas of concentration of the University of Miami Miller School of Medicine (UMMSM). The search strategy was devised by a Calder Medical Library reference librarian, using terms representative of those found in typical searches run by and for Calder patrons. These searches may not be generalized to all searches conducted by researcher-writers because of different subject matter, or because of different library subscriptions to different full text journals.

PubMed was the channel chosen for the searches, since it is the most widely accessible database for medical searching. The most currently available version of QUOSA at the time (8.07.334, build date 09.02.11) was used.

Keywords were used in the searches rather than MeSH terms to insure that the specific words would appear in the results and not be converted into MeSH terms. Results were limited to 40 in order to eliminate randomness that may occur with 30 or fewer results.

Results

The five searches resulted in the retrieval of 159 PDF files and 41 abstracts. Researchers confirmed that QUOSA identified a file as a PDF correctly in every instance. In four cases, retrieved, full text PDFs were identified as only abstracts (no full text) by QUOSA. Full text was available from the publishers' site in two of these instances. The other two full text results that were identified as abstracts were available from an aggregator. The misidentified results

included both PDF and HTML files.

The Performance Review Table shows the terms used in each search and the breakdown of results between PDFs and abstracts for the searches.

When using the QUOSA *Search in Results* feature, QUOSA mistakenly identified results as having all terms present. For example, when searching the terms *macular degeneration antibodies monoclonal*, QUOSA identified 37 results as having all terms present. Manual verification showed only 23 results had all terms. Results in which all terms were highlighted were inconsistent. One search set had no highlighted terms, while another set had all terms highlighted in 30 of 37 results. Since only the first 40 results were retrieved with each search, these inconsistencies cannot be explained by retrieval of older PDF versions.

As indicated in the table, results in which at least one term was not highlighted or was missing altogether was not uncommon.

Discussion

QUOSA is a multifunctional product that shares many features with those reviewed in Part 1 of this article.¹ The major difference between QUOSA and those products is its incorporation of so many features into one program. A user can retrieve, organize, export, and receive alerts about pertinent articles, as well as search within results with terms highlighted for easy identification. QUOSA is not a perfect product, but given its broad functionality, its imperfections seem negligible. Currently, QUOSA is limited in the number of channels with which it integrates. Its performance is strongest with PubMed, making it a product that could be particularly useful for biomedical researchers. Another shortcoming in the version of QUOSA

studied for this paper is that, although QUOSA supports Adobe Acrobat Version 6 and above, some features, such as term highlighting, only work with Version 7 or higher.

Despite these limitations and the minor deficiencies described in the Results section of this paper, QUOSA is a strong and versatile tool for the efficient retrieval, organization and management of data for researcher-writers.

[PLACE TABLE 1 HERE]

Legend: Table 1: QUOSA Performance Review

Conclusion

Research-writers can increase efficiency by using tools to help manage, organize, and search within the full text of articles retrieved. Several useful products are available but downloading them without reading instructions can result in confusion. Most products have introductory videos, which introduce the features of the products. Of all the tools reviewed in Part 1¹, the one that was the most intuitive was Papers. Unfortunately Papers is only available for Macintosh operating systems. QUOSA, although not perfect, appears to be the most robust of the available products. QUOSA Information Manager includes in one program features from several of the products described in Part 1. None of the products reviewed included the full constellation of productivity tools described in the introduction; however these products are rapidly advancing. A time may soon come when a functional, fully integrated product to enhance many aspects of bibliographic research and writing.

[PLACE TABLE 2 HERE]

LEGEND: Table 2: Summary of Features of QUOSA Information Manager

REFERENCES

1. Vaidhyathan, V., Moore, M., Loper, K.A., Van Schaik, J., and Goolabsingh, D. "Making Bibliographic Researchers More Efficient: Tools for Organizing and Downloading PDFs, Part 1: iCyte, Mendeley Desktop, Papers, PDF Stacks, PubGet PaperPlane, WizFolio, and Zotero." *Electron Resour Med Lib*. In press.
2. *About Quosa*. Available: <<http://www.quosa.com/about.htm>>. Accessed: September 10, 2011.
3. *Quosa Information Manager: Quick Reference Guide* (v8.05, PC interface). Available: <<http://calder.med.miami.edu/pages/QUOSA-Quick%20Reference%20Guide-v081215.pdf>>. Accessed: September 8, 2011.
4. *QUOSA Information Manager: User Guide for PC* (8.07). Available: <<http://www.quosa.com/support/helpdocs/pc/807/>>. Accessed: September 12, 2011.
5. *Quosa Help*. Available: <<http://www.quosa.com/support/helpdocs/pc/804/>>. Accessed: September 11, 2011.
6. Moore, M., Van Schaik, J., and Loper, K.A. "Medical self-diagnosis and the web: results from entering plain language symptoms into a search engine." Presented at the Diagnostic Errors in Medicine Annual Conference, Chicago, IL, October 23-26, 2011.

7. Beyrer, C., Baral, S.D., Walker, D., Wirtz, A.L., Johns, B., and Sifakis, F. “The expanding epidemics of HIV Type 1 among men who have sex with men in low- and middle-income countries: Diversity and consistency.” *Epidemiol Rev* 32 (2010): 137-151.
8. Xu, J.J., Henstock, P.V., Dunn, M.C., Smith, A.R., Chabot, J.R., and De Graaf, D. “Cellular imaging predictions of clinical drug-induced liver injury.” *Toxicological Sciences* 105, no.1 (2008): 97-105.
9. Chaussabel, D. “Biomedical literature mining: Challenges and solutions in the “Omics” era.” *Am J Pharmacogenomics* 4 no. 6 (2004): 383-393.