volume. PyscTESTS for example has nearly 30,000 records as opposed to PRQOLID's more than 900.

Although PROQOLID is more limited in scope and content, the quality of PROQOLID records seems to far exceed that of other validated research instrument databases.

BROWSING AND SEARCHING

Users can browse by pathology, population, and author's name. There is also a "Generic Instruments" browse option available for instruments that are not disease-specific, such as the fifteen-dimensional health-related quality of life measure (15D) and the Burden of Treatment Questionnaire (TBQ). Users can also browse an alphabetical list of all available instruments.

Because there is not an overwhelming amount of content in the database, a simple keyword search is usually sufficient. The database also offers some nice-to-have advanced search features, such as the ability to identify instruments available in a specific language and the ability to search for instruments that use a specific mode of data collection (pen and paper, computer administered, etc.). Users can also search for instruments appropriate for a specific population. Available population options include not only age groups and genders, but additional options such as caregivers and terminally ill patients.

LICENSING AND ACCESS

PROQOLID does not offer a free trial period but does offer webinars to those who are interested in obtaining more information prior to subscribing.

Presumably because PROQOL-ID's parent organization, Mapi Research Trust, is a nonprofit organization, their pricing information is freely available on the web in the Mapi Research Trust Online Store. Different subscription rates are available for hospitals and academic institutions, regulatory agencies, and commercial entities. Because Mapi Research Trust is a French organization, their pricing structure is in euros rather than dollars; however, they accept credit card payment and offer billing through subscription agents such as EBSCO Information Services.

The only available access option for this database is username and password. Internet protocol authentication and other access mechanisms are not supported.

OVERALL ASSESSMENT

PROQOLID provides a simple, user-friendly interface and comprehensive information on the instruments it covers. The database's major Achilles heel is its limited content.

Rachel Pinotti, MLIS, AIP, rachel.pinotti@mssm.edu, Gustave L. and Janet W. Levy Library, Icahn School of Medicine at Mount Sinai, New York, NY

REFERENCES

1. US Food and Drug Administration. Clinical outcome assessment qualification program [Internet]. The Administration [cited 21 Jul 2015]. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284077.htm>.
2. Mapi Research Trust. Your instrument in PROQOLID [Internet]. PROQOLID [cited 21 Jul 2015]. <http://www.proqolid.org/authors>.

PubMed2XL (version 2.01). Nitin Arora. http://blog.humaneguitarist.org/projects/pubmed2xl/; free open-source tool for Windows and Linux operating systems.

PURPOSE

PubMed records contain descriptive metadata such as author, abstract, subject headings, and grant numbers. Getting the metadata into a spreadsheet program like Microsoft Excel or Open Office Calc allows users to sort, filter, and transform the data for new purposes. Once the data are in a spreadsheet, bibliometric analysis is possible, as is creating charts and figures. Combining the spreadsheet with a word processor mail merge function enables a bibliography to be quickly generated. Despite the benefits of having PubMed metadata in a spreadsheet, PubMed's built-in spreadsheet functionality—the comma separated value (CSV) download—is severely lacking. Thankfully, an open-source tool called PubMed2XL meets this needed functionality.

DESCRIPTION

PubMed2XL is a Windows or Linux application that converts PubMed extensible markup language (XML) into Microsoft Excel or OpenDocument spreadsheets [1]. The program is freely available under an MIT license [2]. PubMed2XL uses the Python programming language and converts

Spreadsheet column name	Example	Notes
PMID	23023988	Hyperlink to PubMed record
Digital object ID	10.1002/pds.3334	
Article Title	The incident user design in comparative effectiveness research	
Abstract	Comparative effectiveness research includes cohort studies and registries of interventions. When investigators design…	
Language	eng	
Journal: Title	Pharmacoepidemiology and drug safety	
Journal: Abbreviation	Pharmacoepidemiol Drug Saf	
Journal: NLM ID	9208369	Hyperlink to NLM catalog record
Journal: Year	2013	
Journal: Month	Jan	
Journal: MEDLINE Date	2015 May–Jun	
Journal: Volume	22	
Journal: Issue	21	
Journal: MEDLINE Pagination	1–6	
Citation Medium	Internet	
First Author (First Last)	Eric S Johnson	
All Authors (Last, Initials)	Johnson, ES; Bartman, BA; Briesacher, BA; Fleming, NS; Gerhard, T; Kornegay, CJ; Nourjah, P; Sauer, B; Schumock, GT; Sedrakyan, A; Stürmer, T; West, SL; Schneeweiss, S	
Affiliation	The Center for Health Research, Kaiser Permanente, Portland, Oregon, USA. Eric.S.Johnson@kpchr.org	
Related Citations	click here	Hyperlink to PubMed similar articles in summary view
Related Citations (XML format)	click here	Hyperlink to PubMed similar articles in XML view
Related Articles by Review	click here	Hyperlink to subset of PubMed similar articles, with publication type Review, in summary view
Related Articles by Review (MEDLINE format)	click here	Hyperlink to subset of PubMed similar articles, with publication type Review, in MEDLINE view
Article Type	journal	

Table 1

Spreadsheet columns created by the PubMed2XL default stylesheet

PubMed XML records to a spreadsheet via EXtensible Stylesheet Language Transformations (XSLT) [3]. PubMed2XL was written by Nitin Arora and includes software developed by Roman V. Kiseloiv. The current version is 2.01.

AUDIENCE

The audience for this resource is any PubMed user who wants to work with metadata from search results in a spreadsheet program, such as librarians, researchers, clinicians, and administrators. The tool is especially helpful for digital initiative or repository librarians who need to transform PubMed data for use in other systems.

MAJOR FEATURES

As an input, PubMed2XL takes a downloaded PubMed XML file and then outputs either a Microsoft Excel (.xls) or Open Office (.ods) file. Within seconds, the spreadsheet is automatically saved to the same directory as the source file, and the program offers to open the spreadsheet. The user can choose whether to include citations for books in the spreadsheet results.

To create the spreadsheet, the program uses another XML file called a stylesheet. The stylesheet controls the column titles, their order, and what data are included in the cells. PubMed2XL comes with two stylesheets. One, aptly named pmid_only.xml, creates a spreadsheet with a single column listing the PubMed Identifiers (PMIDs). The other stylesheet, called default.xml, creates a spreadsheet with twenty-three columns (Table 1).

One of the most useful features of PubMed2XL is that it allows users to create their own stylesheets. Using the default stylesheet as a starting point, I created one to capture additional metadata like Medical Subject Headings (MeSH) terms, grant numbers, and International Standard Serial Numbers (ISSNs). I also turned the PMC and DOI data into hyperlinks in my spreadsheet (the default stylesheet already does this with the PMID). Finally, using XSLT markup, I automated a few tasks to clean up the data entered into the spreadsheet. For

example, my stylesheet combines the various XML date nodes into a single spreadsheet column in the YYYY-MMM-DD format.

USABILITY

PubMed2XL is easy to use on the Windows platform; I did not test it on Linux. The program comes with an installer for Windows, making installation quick and routine. The program is relatively small (59 MB), so it can be run off most USB flash drives—an advantage if you want to test the program before installing it on your main computer. The narrow scope of the program makes for simple graphic user interface (GUI) navigation as there are only a few options in each menu.

DOCUMENTATION AND SUPPORT

The documentation for the program is kept on the creator's website [1]. Arora, the creator of PubMed2XL, also has a five-minute video explaining installation and use [4]. The video, which is four years old, was recorded for the first version of PubMed2XL. Though the user interface and default stylesheet have changed in the intervening years, the video still provides a good overview of how to use the program.

Initially, I had trouble using XSLT to get MeSH descriptors and qualifiers to appear in the same spreadsheet column. I reached out to Arora via the website and received a response within 24 hours. Since the release of version 0.9 in 2010, Arora has been engaging with users via the comment section on the website.

LIMITATIONS

The main limitation of the program is the number of PubMed records that can be processed at a time. The creator of the program recommends that input files contain fewer than 5,000 records, but the creator was able to process 25,000 records from the command line using a development version of the software [1]. Given that it takes less than a minute to process 5,000 records, the recommended limit has not been a problem for me; it is simple enough to divide a large PubMed result set into smaller XML downloads. A quibble is that there have been no updates to the program since October 2013 [5], but it is not clear that a new version is needed, as the program is a stable release and works well. Finally, creating a new stylesheet does require experience with XSLT, so a user unfamiliar with XML may find it difficult to create a new stylesheet from scratch.

COMPARISON

PubMed itself offers a CSV download feature to open search results in a spreadsheet, but the resulting spreadsheet is far inferior to the one created using PubMed2XL's default stylesheet. PubMed's CSV download has several drawbacks. First, it includes only a few metadata fields; metadata like abstract and MeSH terms are not available. Second, the resulting spreadsheet columns have ambiguous labels. For example, a column labeled "Description" actually contains author information, and "Properties" includes the create date and first author. Third, some of the columns have limited value: "Resource" always contains data saying "PubMed," while "DB" consistently lists ''pubmed,'' and "Type" contains "citation." Finally, the CSV download feature in PubMed cannot be customized by the end user.

CONCLUSION

PubMed2XL is an easy-to-use program that transforms PubMed XML data into a spreadsheet. The default stylesheet provides metadata in a better format than the CSV download that PubMed offers does. Librarians can use the program to easily extract data from PubMed and prepare it for other systems such as institutional repositories, although creating a new stylesheet with PubMed2XL requires knowledge of XSLT to unlock the full potential of the program.

David Isaak, MSLIS, david.c.isaak@kp.org, Kaiser Permanente Center for Health Research, Portland, OR

REFERENCES

1. Arora N. PubMed2XL [Internet]. Nitin Arora; 2010 [cited 8 Aug 2015]. <http://blog.humaneguitarist.org/projects/pubmed2xl/>.

2. Open Source Initiative. The MIT license (MIT) [Internet]. The Initiative [cited 8 Aug 2015]. <http://opensource.org/licenses/MIT>.

3. World Wide Web Consortium. Transformation [Internet]. The Consortium [cited 8 Aug 2015]. <http://www.w3.org/standards/xml/transformation>.

4. Arora N. PubMed2XL: basic installation and use [Internet]. Nitin Arora; 2010 [cited 8 Aug 2015]. <https://vimeo.com/15098984>.

5. Arora N. PubMed2XL 2.01 available [Internet]. Nitin Arora; 5 Oct 2013 [cited 8 Aug 2015]. <http://blog.humaneguitarist.org/2013/10/05/pubmed2xl-2-01-available/>.