Reference Management meets Web 2.0

Martin Fenner, MD

Carl-Neuberg-Str. 1, Hannover Medical School, 30625 Hannover, Germany

Correspondence: Martin Fenner, Department of Hematology, Hemostaseology, Oncology and
Stem Cell Transplantation, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany.
Phone +49 511 532-4077, Fax +49 511 532-8077, Email: fenner.martin@mh-hannover.de

Abstract

Reference management software has been used by researchers for more than 20 years to find, store, and organize references, and to write scholarly papers. Recently developed collaborative web-based tools have resulted in a number of interesting new features, and in a number of new reference managers. These developments are changing which reference managers we use, and how we use them.

Keywords: reference management, Web 2.0, citation

Introduction

Reference management software allows researchers to create a personal electronic collection of relevant scholarly publications, and to use this collection to write their own scholarly works. Some of the earliest programs (e.g. Endnote and BibTeX) have now been around for 25 years, and their core functions remain unchanged. The first big change came around 15 years ago when references were no longer typed in by hand, but rather retrieved from the Internet. Another important change started five years ago in the context of what is typically called Web 2.0. Reference managers no longer were used as software on a single desktop computer. They now more often than not are web-based applications (with or without synchronization to a desktop version), and this allows easy sharing of references between several users, computers and/or other web applications. This trend has led to a number of new reference managers, both commercial and freely available.

Reference managers help researchers by performing three basic functions:
1) **Searching**: find relevant scientific literature,
2) **Storing**: store the results of that search in a personal database for later retrieval, and
3) **Writing**: insert references when writing a manuscript.

Although all three functions could also be performed without specialized software, a manual approach is not recommended for managing anything beyond a handful of references.

Managing references manually is much slower and prone to errors, e.g., when renumbering all references in a manuscript after inserting a new citation.

In the next sections, we will look at the three basic reference manager functions in more detail, with special emphasis on how the newer Web-based reference managers enhance some of these functions with "social" features.

Searching

All reference managers can import references from bibliographic databases, either by directly searching these databases (PubMed, Scopus, Web of Science, Google Scholar, etc.), and/or via so-called bookmarklets. Bookmarklets are specialized bookmarks for web browsers that retrieve references from web pages, e.g., the result of a PubMed search. Both strategies have their advantages, but for most users there is no real difference. Interesting references are also showing up in other places besides bibliographic databases. Most commonly this is a journal table of contents received via email or RSS reader, but it can also be a blog post or even a Twitter message. In order to retrieve this reference information, users usually first have to follow a link to a bibliographic database or journal web page.

A traditional search strategy typically uses keywords, author and journal names, and publication dates. Because web-based reference managers such as CiteULike or Mendeley store millions of references by thousands of users, they also offer a very powerful "social" search. They can show you the references of users with similar interests, or papers similar to the paper you just imported. This "social" search requires a critical mass of users, but might in a few years surpass traditional search strategies in popularity. Sharing references in private or public groups is already a very popular feature of web-based reference managers. Reading lists — lists of references required for a particular course — are one typical use.

Storing

Reference managers are databases that store references. In the life sciences a reference is typically a journal article, and sometimes a conference abstract, book chapter, or web page. But reference managers can also handle a long list of other references. Among the 48 reference types supported in the latest version of Endnote (Endnote X4), some of the lesser known are online database, audiovisual material, grant, blog, and research dataset. This wider definition of a reference looks not unlike the bookmarks we store with our web browser. Bookmarks can be stored in specialized websites (e.g., delicious), and several of these so-called social bookmarking sites also handle scientific references (CiteULike being the most popular).

Since almost all scholarly publications are now published in electronic form, reference managers (most notably Papers) a few years ago started to not only manage references, but also store the fulltext PDF files associated with them. This is the most convenient way to store these PDF files. And it has another advantage: we can do powerful fulltext searches of the publications stored in our reference manager. Many reference managers can import PDF files and extract the reference information from the PDF file. And some of them have an integrated PDF viewer that allows highlighting of text and note taking.

As reference managers are basically databases, they should allow the user to import and export references (RIS is the best standard file format), find duplicate records, or group references by subject or keyword. Only some reference managers can also group references by author or journal, or list all references cited by a particular reference.

Many reference managers offer a web-based version. This allows users to have the same reference database on more than one computer and to share references with others. Whe-

Reference Manager Overview

compiled by Martin Fenner
Version 2.4, September 19, 2010
http://bit.ly/refman

	Zotero	EndNote	Mendeley	Refworks	CiteULike	JabRef	Papers	Citavi
Search — Search bibliographic databases								
Pubmed	■	■		■		■	■	■
Scopus				■			■	■
Web of Science				■			■	■
Bookmarklet	■		■	■	■			■
Share — Store online, sync, share with others								
WWW	■	■	■	■	■			
PDF files	■		■	■	■			
Public folders	■		■	■	■			
API	■	■	■		■			
Store — Organize references on computer								
Windows	■	■	■	■		■		■
Macintosh	■	■	■			■	■	
Linux	■		■			■		
iPhone/iPad			■				■	
Read — Import, search, read and organize PDF files								
Extract metadata	■	■	■			■	■	
Fulltext search	■	■	■	■			■	
PDF viewer			■				■	
File organizer			■				■	
Write — Put formatted references into manuscript								
Microsoft Word	■	■	■	■				■
Open Office	■	■	■					
LaTex			■			■		
Edit styles	■	■		■				■

Table 1. Some popular reference managers with Web 2.0 features

reas some reference managers (e.g. CiteULike, RefWorks) are only web-based, others (e.g., Mendeley, Endnote, Zotero) synchronize a desktop with a web version. This year we have seen a proliferation of reference management tools for mobile devices such as the iPhone, and also the first reference managers for the iPad.

Writing

Reference managers are a big time saver in manuscript writing. They help in inserting citations into the text and automatically create a bibliography in the desired citation style. Although many reference managers now come with more than 1000 different citation styles, only a few of them (e.g., Endnote or Refworks) allow the user to edit them — an important feature for some users. Not all reference managers have a word processor plugin, and if they do they often only support Microsoft Word and maybe OpenOffice. The built-in reference management features of the latest version of Microsoft Word (Word 2007 or Word 2008 for Macintosh) are very rudimentary and not recommended.

Almost all scientific papers are now written by more than one author. Collaborative online writing tools such as Google Docs or Zoho Writer facilitate the writing process, as authors don't have to repeatedly send around draft versions of manuscripts via email. Unfortunately no reference manager directly supports these online tools beyond a simple copy and paste.

Conclusions

Reference management software is currently undergoing a lot of exciting changes, and now is a good time to test some of the programs mentioned in this article. Several of the newer reference managers are free to use, so cost shouldn't be a reason not to start using such a tool. CiteULike and other web-based tools can be used right away without installing any software; downloading and installing one of the free tools (e.g. Zotero or Mendeley) also takes less than an hour. Every reference manager has strengths and weaknesses; the comparison chart can help with finding the right tool to get started. It is also a good idea to start with a reference manager that your colleagues use. Not only can they help you with questions, but you also want to use the same reference manager when writing a manuscript together.

Acknowledgements

Links

Connotea, http://www.connotea.org/
Endnote/Endnote Web, http://www.endnote.com/
Refworks, http://www.refworks.com/
Zotero, http://www.zotero.org/
Mendeley, http://www.mendeley.com/
CiteULike, http://www.citeulike.org/
Jabref, http://jabref.sourceforge.net/
Papers, http://mekentosj.com/papers/
Citavi, http://www.citavi.com/en/index.html

References

1. Hull D. et al. Defrosting the digital library: bibliographic tools for the next generation web. PLoS Comput Biol. 2008;4(10):pp.e1000204 http://www.ploscompbiol.org/article/info:doi/10.1371/journal.pcbi.1000204

2. Innovations in Reference Management Workshop, http://www.open.ac.uk/telstar/event

3. Reference Manager Overview, http://bit.ly/refman