Docear4Word: Reference Management for Microsoft Word based on BibTeX and the Citation Style Language (CSL)

Joeran Beel Docear, Magdeburg Germany

Stefan Langer Docear, Magdeburg Germanv

Marcel Genzmehr Docear, Magdeburg Germany

beel@docear.org

langer@docear.org

genzmehr@docear.org

ABSTRACT

In this demo-paper we introduce Docear4Word. Docear4Word enables researchers to insert and format their references and bibliographies in Microsoft Word, based on BibTeX and the Citation Style Language (CSL). Docear4Word features over 1,700 citation styles (Harvard, IEEE, ACM, etc.), is published as open source tool on http://docear.org, and runs with Microsoft Word 2002 and later on Windows XP and later. Docear4Word is similar to the MS-Word add-ons that reference managers like Endnote, Zotero, or Citavi offer with the difference that it is being developed to work with the de-facto standard BibTeX and hence to work with almost any reference manager.

Categories and Subject Descriptors

H.4.1 [Information Systems Applications]: Office Automation word processing, workflow management.

General Terms

Management, Documentation

Keywords

bibtex, citation style language, CSL, microsoft word, reference management, reference manager

1. INTRODUCTION

Reference management is probably the most tiring task for students and researchers. They have to re-type and format bibliographic information over and over, for each paper, assignment or thesis. This is particularly frustrating if they need to change citation styles in a document. This might become necessary, for instance, because a supervisor changes his mind on his favorite citation style, or a paper is submitted to another journal which requires a different citation style than the previous journal.

In the past decades many software tools evolved to facilitate this workflow. Commercial tools such as Endnote and Citavi enable researchers to maintain a database with all the bibliographic data of their references. These so called 'reference managers' usually offer add-ons for Microsoft Word allowing users to insert and format references and bibliographies in a convenient way. Also some open source tools offer such add-ons, for instance Zotero. However, all these tools use proprietary data formats. Accordingly, a Microsoft Word add-on from one of these tools (e.g. Endnote) works only with the corresponding data format of that tool.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

JCDL'13, July 22-26, 2013, Indianapolis, Indiana, USA. ACM 978-1-4503-2077-1/13/07.

There is one alternative to the proprietary formats, namely BibTeX. BibTeX was created by Oren Patashnik in 1988 [4] and is the defacto standard format in reference management. There are many reference mangers directly supporting BibTeX, for instance JabRef, BibDesk and our own reference manager Docear [1]. Even proprietary tools such as Endnote usually allow exporting their database to BibTeX. There is a Microsoft Word Add-On for BibTeX based databases named BibTeX4Word¹. However, BibTeX4Word requires the installation of additional tools and is difficult to setup and use. For instance, in the Blog MedicalNerds more than 250 comments were made on BibTeX4Word, mostly questioning about the usage - especially installing new citation styles is complicated [3].



Fig. 1. Docear4Word ribbon in Microsoft Word 2010

We developed Docear4Word², a Microsoft Word add-on to insert and format references directly in MS Word based on BibTeX. Docear4Word is open source and runs with Microsoft Word 2002 and later on Windows XP and later. After the installation, Docear4Word is accessible in the "Reference" ribbon if Microsoft Word 2007 or later is used (Fig. 1). In Word 2002 and 2003 a separate toolbar is installed. Docear4Word was primarily intended for users of our literature management tool Docear [1], but can be used with any BibTeX file from any reference manager. In contrast to BibTeX4Word, Docear4Word is more user-friendly and uses the citation style language (more details in the following section).

The remainder of this paper provides a detailed overview of Docear4Word.

2. DOCEAR4WORD

2.1 Maintaining a BibTeX database

BibTeX is a text-based format. Accordingly, a BibTeX file can be created and edited with any text editor. However, there are several tools offering a graphical user interface to create BibTeX files, for instance JabRef and our own reference manager Docear. Fig. 2 shows a screenshot of Docear. Docear provides a graphical user interface for specifying title, authors, and other bibliographic data of academic literature. Based on this data Docear automatically creates a BibTeX entry. Instead of Docear, or a text editor, any reference

¹ http://www.ee.ic.ac.uk/hp/staff/dmb/perl/index.html

² http://www.docear.org/software/add-ons/docear4word/overview/

management tool can be used that uses the BibTeX format or that may export its proprietary format to BibTeX.

2.2 Inserting references in Microsoft Word

Fig. 3 shows the dialog to search and insert references. The dialog allows selecting several references at once and specifying individual page numbers. Once, the references are selected and the "Add References" button is clicked, references are added in the document and formatted accordingly to the selected citation style (Fig. 5).

Pefault SansSerit 10 B Image: SansSerit 10 B Image: SansSerit Image: Sa	ing F <u>i</u> lter	<u>T</u> ools <u>M</u> ap:	s <u>H</u> elp						
Author Author Title Author Title Article Article Article Article Article Article Article Article Article Author Title Author Title Article Author Title Article Artic	🔨 🥕 Default 💽 SansSerif 💽 10 💽 🖪 🖌 👁 👷 🔧 🥆 🏵 🤶								
aferenceDB.bib Entryt Author Title Bakkalbasi et (D) Article Bjart-[Njan (M) Ich Index? - {A} comparison of (M)o(S), (S)copus and (C) (Article Bjart-[Njan (M) Ich Index? - {A} comparison of (M)o(S), (S)copus and (C) (D) Inproc Bjeel and (G)ipp (G) cogle (S)cholar's (R)anking (A)lgorithm: (The (I)mpact of (C) (D) Inproc Bjeel and (G)ipp (G) cogle (S)cholar's (R)anking (A)lgorithm: (The (I)mpact of (C) (D) Inproc Bjeel and (G)ipp (G) cogle (S)cholar's (R)anking (A)lgorithm: (The (I)mpact of (C) (D) Inproc Bjeel and (G)ipp (G) cogle (S)cholar's (R)anking (A)lgorithm: (The (I)mpact of (C) (D) Inproc Bjeel and (G)ipp (Article Bjeel and (G)ipp (Article Bjeel and (B)ipp (J) coran (B)eel and (B)ela General Author (J) oeran (B)eel and (B)ela (J) oeran (B)eel and (B)ela G)ipp Title (A)cademic search engine spam and (G)cogle (S)cholar's resilience against it Journal Journal of Electronic Publishing Year 2010 Maturee	attributes								
Entryt Author Title Article (B)akkalbasi et (T)hree options for citation tracking: (G)oogle (S)cholar, (S)copus and (G)active (B)ar-(I)lan Article (B)ar-(I)lan (W)hich h-index? - (A) comparison of (W)o(S), (S)copus and (G)ipp (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)inproc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)inproc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)inproc (B)eel and (G)ipp (D) Inproc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)inproc (B)eel and (G)ipp (O) the (R)obustness of (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)inproc (B)eel and (G)ipp (O) the (R)obustness of (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)inproc (Aticle (B)eel and (G)ipp (A)cademic Search engine Spam and (G)oogle (S)cholar's resilience against it (J)oeran (B)eel and (B)ela (G)ipp (J)oeran (B)eel and (B)ela (G)ipp Title (A)cademic search engine spam and (G)oogle (S)cholar's resilience against it Journal Journal of Electronic Publishing Year 2010	+ 🗾 🖃	II 🕅	🥖 🖉		🗅 💾 💾				
Article (B)akkalbasi et (T)hree options for citation tracking: (G)oogle (S)cholar, (S)cop Article (B)ar-{[I]an (W)hich h-index? - (A) comparison of (W)o(S), (S)copus and (C Improc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)gorithm: (T)he (I)mpact of (/ Improc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)gorithm: (T)he (I)mpact of (/ Improc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)gorithm: (T)he (I)mpact of (/ Improc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)gorithm: (T)he (I)mpact of (/ Improc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)gorithm: (T)he (I)mpact of (/ Improc (B)eel and (G)ipp (O)actemic search engine spam and (G)oogle (S)cholar's resilience (D)actemic (D)	eferenceDB	l.bib							
Inproc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)mproc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)mproc (B)eel and (G)ipp (D) Inproc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)mproc (B)eel and (G)ipp (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mpact of (C)mproc (B)eel and (G)ipp (D) Inproc (B)eel and (G)ipp (G) the (R)obustness of (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mproc (B)eel and (G)ipp (O) the (R)obustness of (G)oogle (S)cholar's (R)anking (A)lgorithm: (T)he (I)mproc (B)eel and (G)ipp (A)cademic search engine spam and (G)oogle (S)cholar's resilience (S)cholar's resilience (S)cholar's resilience against it (J)oeran (B)eel and (B)ela (G)ipp (A)cademic search engine spam and (G)oogle (S)cholar's resilience against it Journal Journal of Electronic Publishing Year 2010		- ,			e options for o	citation tracking	g: {G}oogle {S	}cholar, {S}cop	
Improc (B)eel and (G)ipp (O) the (R)obustness of (G)oogle (S)cholar (A)gainst (S)pam Article (B)eel and (G)ipp (A)cademic search engine spam and (G)oogle (S)cholar's resilience (A)ECOU-(C) bits Required fields Optional fields General Abstract Review BibTeX sourc Author (J)oeran (B)eel and (B)ela (G)ipp General Abstract Review BibTeX sourc Author (J)oeran (B)eel and (B)ela (G)ipp General Abstract Review BibTeX sourc Author (J)oeran (B)eel and (B)ela (G)ipp General Abstract Review BibTeX sourc Author (J)oeran (B)eel and (B)ela (G)ipp General Abstract Review BibTeX sourc Journal Journal of Electronic Publishing Journal of Electronic Publishing Journal of Electronic Publishing Journal of Electronic Publishing Year 2010 Journal Journal of Electronic Publishing Journal of Electronic Publishing	🍌 🖸 Inp	oroc {B}e	el and {G}ipp el and {G}ipp	{G}000 {G}000	gle {S}cholar's gle {S}cholar's	{R}anking {A} {R}anking {A}	lgorithm: {T}h lgorithm: {A}r	ne {I}mpact of {({}}ntroductory {	
Required fields Optional fields General Abstract Review BibTeX source Author {Joeran {B}eel and {B}ela {G}ipp Image: Comparison of the second of the secon	np 🖸 🙏	Display [] </td							
Image: Title [J]oeran (B)ela (G)ipp Title [A]cademic search engine spam and (G)oogle (S)cholar's resilience against it Journal Journal of Electronic Publishing Year 2010									
Algo addemic search engine spam and (G)oogle (S)cholar's resilience against it Journal Journal of Electronic Publishing Year 2010	Author	{J}oeran {B}eel and {B}ela {G}ipp							
Year 2010	Title	{A}cademic search engine spam and {G}oogle {S}cholar's resilience against it							
2010	Journal	Journal of Electronic Publishing							
	Year	2010							
		[

Fig. 2. Maintaining the BibTeX database with an external reference manager (e.g. Docear)

Pages	Bib Tex Key	Title 🔺	Authors	Year	Timestamp
180	Beel10	Academic Search Engine Optimization (ASEO): Op	Joeran Beel and Bela Gi	2010	2009-09-04
	Beel2010	Academic search engine spam and Google Schola	Joeran Beel and Bela Gipp	2010	2010-12-17
	Beel09	Google Scholar's Ranking Algorithm: An Introduct	Joeran Beel and Bela Gipp	2009	2009-02-24
	Beel09b	Google Scholar's Ranking Algorithm: The Impact o	Joeran Beel and Bela Gipp	2009	2009-02-24
	Beel09a	Google Scholar's Ranking Algorithm: The Impact o	Joeran Beel and Bela Gipp	2009	2009-02-24
	Gipp10a	Integrating Citation Proximity Analysis into Google	Bela Gipp and Joeran Beel	2010	2010-02-01
▼	Beel10c	On the Robustness of Google Scholar Against Spam	Joeran Beel and Bela Gipp	2010	2010-04-21

Fig. 3. Selecting and inserting a reference

2.3 Changing the Citation Style

Docear4Word uses the citation style language [5] and the citeproc-js citation processor [2] to format references. The citation style language supports more than 1,700 citation styles such as IEEE, Harvard, MLA and ACM in several variations. Docear4Word users can select a style in the style box (Fig. 4) and install new styles from the style repository. When a new style is selected all references are formatted accordingly.

2.4 Insert a Bibliography

Docear4Word automatically creates a bibliography based on the references in the body of the document (Fig. 5). The user can

choose where to insert the bibliography and the bibliography is automatically updated when new references are inserted.

Page Layout References Mailing	gs Review View	Add	-Ins
Style: Insert a citation	+ Add Reference	Style:	ACM SIG Proceedings
insert bibliography	🧮 Insert Bibliography		Find more Citation Styles
tion Refresh	Refresh		ACM SIG Proceedings With Long Author List
Mendeley Cite-O-Matic			ACM SIG Proceedings
	ACS Chemical Biology		
1 . 1 2 3 4 5 .	Advanced Engineering Materials		
		American Psychological Association 6th Editio	
Recently, we published some pap	American Psychological Association 5th Editio		
researched Google Scholar [1, 5,	Chicago Manual of Style (note, annotated bi		
	Frology		

Fig. 4. Style chooser

Page Layout References	Mailing	s Review View	Add-Ins					
Style: Insert a citation *		+ Add Reference	Style: ACM SIG Proceedings *					
		📑 Insert Bibliography	A Previous Vext	Citations &	Capti			
Insert Citation 😰 Refresh	Export	C Refresh	🕒 About 🄛 Settings	Bibliography *	-capti *			
Mendeley Cite-O-Matic			Docear4Word					
······································								

Recently, we published some papers about Google Scholar [2–4]. There are also many other authors who researched Google Scholar [$1_{4,5,6}$].

References

- Bar-Ilan, J. 2007. Which h-index? A comparison of WoS, Scopus and Google Scholar. Scientometrics. 74, (2007), 257–271.
- Beel, J. et al. 2010. Academic Search Engine Optimization (ASEO): Optimizing Scholarly Literature for Google Scholar and Co. Journal of Scholarly Publishing. 41, (Jan. 2010), 176–190.
 Beel, J. and Gipp, B. 2009. Google Scholar's Ranking Algorithm: The Impact of Articles' Age (An

] Beel, J. and Gipp, B. 2009. Google Scholar's Ranking Algorithm: The Impact of Articles' Age (An Empirical Study). Proceedings of the 6th international Conference on Information Technology: New Generations (ITNG '09) (Las Vegas (USA), Apr. 2009), 160–164.

Fig. 5. Automatically created bibliography

3. OUTLOOK

Docear4Word was released as final and stable version 1.0 on http://docear.org. However, we will continue to improve Docear4Word. Among others, it is planned to offer a version for Microsoft Word on MacOS; implement support for footnotes; enable suppressing author and/or year in a reference; implement an installer for new citation styles; and allow using multiple BibTeX files at the same time. As Docear4Word is available as open source, we sincerely invite other researchers to join the development.

4. ACKNOWLEDGEMENTS

We are very grateful for the Citation Style Language and CiteProc-JS by Bruce D'Arcus, Frank G. Bennett, Jr. Rintze M. Zelle, and many others. In addition, we are glad for all the support we received from Simon Hewitt.

5. REFERENCES

- Beel, J., Gipp, B., Langer, S. and Genzmehr, M. 2011. Docear: An Academic Literature Suite for Searching, Organizing and Creating Academic Literature. *Proceedings* of the 11th JCDL (2011), 465–466.
- [2] Frank G. Bennett, J. 2011. The citeproc-js Citation Processor. http://gsl-nagoya-u.net/http/pub/citeproc-doc.html
- [3] Medicalnerds 2007. How to use JabRef (BibTeX) with Microsoft Word 2003. http://www.medicalnerds.com/how-touse-jabrefbibtex-with-microsoft-word-2003/
- [4] Patashnik, O. 1988. BibTexing. http://bibtexml.sourceforge.net/btxdoc.pdf
- [5] Zelle, R.M. 2010. Citation Style Language 1.0. http://citationstyles.org/2010/03/22/citation-style-language-1-0/