

# Bib Manager and Word Citer: Bibliography Management and Citation Extraction

Fernando Sáenz-Pérez

Website <http://www.fdi.ucm.es/profesor/fernan>

Address Facultad de Informática, Universidad Complutense de Madrid, Spain

**Abstract** This article describes two tools for managing and using BibTeX bibliographies. The first tool, Bib Manager, allows the user to manage references and store them both as relational databases as well as the usual .bib text plain files. The second tool, Word Citer, allows the user to cite and list references, stored in a database or in a .bib file, in a MS Word document with the selected bibliography style. It relies on another open-source project, LaTeX2RTF [7], which we use to provide format to the list of references as well as generating the citation text. Bib Manager is cross-platform, free, and open-source. Word Citer has the same features but, obviously, it is not cross-platform. Both of them will be distributed under GPL.

*Keywords:* Tools, Free, Open-source, Cross-platform, BibTeX, Bibliography Management, Databases

## 1 Introduction

Our role in the university as researchers (hence, paper-writers) and lecturers (thus, manual and study-guide makers) requires us to deal with a few to numerous references in our documents. It is not uncommon to have BibTeX files of around 1MB. Several of us often use both LaTeX and MS Word documents depending on the document we are preparing. Sometimes, we use MS Word for rapid preparation, but when we need either professional-looking documents, or we are preparing large documents, or including lots of references, we resort to LaTeX and BibTeX. BibTeX provides versatile bibliography handling, and makes the listing of references an easy task. However, since references are stored as plain tex

Copyright © 2007 Fernando Senz-Prez.

Permission is granted to distribute verbatim or modified copies of this document provided this notice remains intact.

files, when the number of references grows the task of managing them becomes more difficult. Therefore, we were looking for, first, a bibliography manager, and, second, a reference citer for MS Word which could use those bibliographies. I directed a team of students in developing such tools under my guidance. As a result, a project was started as part of their university studies to develop a bibliography manager as well as a citer for Word documents. The requirements of this project were several-fold:

*Database-oriented* A bibliographic (data) file is usually present but bibliographic databases are often missing. Databases, however, add several features which cannot be found in file-oriented applications. The most important feature for us is the concurrent access to databases. This allows concurrent updating by different users, who can seamlessly add references to a common bibliographic database. In fact, this is one of our current problems in writing collaborative documents, where each writer in the team has to add references. Therefore, using a database greatly simplifies our work. We focused on relational databases because they are quite commonly used nowadays, and their capabilities fulfill our requirements.

*Simple* We looked for, first, a simple, intuitive GUI. Second, we wanted to have the option of working with a database or simply with `BIBTEX` plain files. Third, the tool should be easily installed, possibly requiring no installer. In fact, merely decompressing an archive and executing a file is enough to start working with Bib Manager. Unfortunately, in the current version of Word Citer we were not able to provide a simple installer; it requires a manual and cumbersome installation procedure. Obviously, this will be fixed as soon as possible.

*Internationalized* In order to have a wider audience, text files are used to define all the texts of the GUI, so that the system can be easily localized to different languages.

*Open-source* We do believe that the open-source alternative will produce better software and systems, and allow developers to contribute enhancements. This is fitting in a university context, so that students will have sources available to learn and practice with them. In an unusual step, the project's main application has been developed without using other software as a basis. While doing this would

have allowed more rapid development, starting from scratch gave the students a unique opportunity, one of the few they will probably have in their careers. They were able to rely on L<sup>A</sup>T<sub>E</sub>X distributions for lists of references and cross-translations between B<sub>I</sub>B<sub>T</sub>E<sub>X</sub> and database entries. In addition, they used the free, open-source project LaTeX2RTF [7].

*A Cross-Platform System* Despite the basic need to develop the same system for several platforms, many of us use several different operating systems and therefore want to have the same applications available on several platforms. As a consequence, the selected development language was Java, which is widely known and works on any platform. Of course, this was not the only alternative, since one can use C#, for example. But for developing a cross-platform system that can be widely used and modified in an open-source project, Java was chosen.

*Free* One of our goals in the university is to share knowledge, and we make our work freely available to others. Further, providing free software allows for widespread usage. In addition, our university and several national projects support us as researchers and lecturers (see Section 5), so we do not need additional income from developing the software.

Before starting this project, other related systems were analyzed. Note that there are many bibliography managers and we only list here some representative ones. On the one hand, there exist some commercial or non-free database managers such as P<sub>A</sub>P<sub>Y</sub>R<sub>U</sub>S (Windows and Macintosh, which now becomes free but not open-source) [9], ProCite (Windows, personal and networked) [10], Reference Manager (Windows and Macintosh) [11], EndNote (Windows and soon for Macintosh) [5], and RefWorks (a Web-based application to create personal bibliography from text files or online databases) [12]. We reviewed their features but did not consider them an alternative to free software.

On the other hand, among free software systems, we found SIXPACK [13], JabRef [6], and BibShare [3]. SIXPACK is open-source, cross-platform and manages file-oriented bibliographies. In addition, it does not make automatic citations or lists of references, and it seems to be a closed project (no Web page modifications since 2000). JabRef is open-source and cross-platform. It is also file-oriented and does not include a citer. BibShare (an evolution of BibWord [4]) is free but not

open-source (although it is open in the sense that it can be connected to different word processors by programming the interface), and only works for Windows.

In summary, after stating our original goals and reviewing related systems, we concluded that no system fully fits our requirements.  $\LaTeX$  and/or MS Word users might find the tools we propose useful. Note, however, that our system is emerging and it is in the early stages of development. However its future development will be guided by the aforementioned objectives as well as user feedback.

This paper is concisely organized as follows. Sections 2 and 3 describe Bib Manager and Word Citer, respectively, from a user point-of-view. Some conclusions are summarized and some future work is pointed out in Section 4. Finally, acknowledgements are posed in Section 5.

## 2 Bib Manager

### 2.1 Features and Limitations

This tool is basically a bibliographic manager based on the  $\BibTeX$  data format. Data are stored in databases, instead of files, so that concurrent access from different (remote) instances of Bib Manager is possible, therefore maintaining a consistent storing for shared bibliographies. MySQL and MS Access database management systems (DBMS) are supported, and others will also be supported soon (PostgreSQL, Oracle, Sybase, DB2, etc.). Note that although Access is a personal, file-oriented, passive (without triggers) database, it can be remotely accessed with a networked file or as a Web service. In addition, locking and transactions are supported through their APIs, so that concurrent accesses can be ensured to be isolated. Several databases can be opened in Bib Manager at a time. It can import  $\BibTeX$  data from `.bib` files and export a database to a `.bib` file. Entries in a database can be created, modified, deleted and copied to any other open database. A key for a new reference can be automatically created following the alpha bibliographic style.  $\BibTeX$  display is always possible either for single references or for the complete database. (Relational) table views can be filtered by a search filter for usual fields and for any fields. The table view can be ordered by any column, columns can be hidden or shown, and they can be moved and sized. For the modification of a single entry, a dialog is provided; in the table

view, direct modifications will be possible. A log is provided for development purposes.

As far as limitations, Bib Manager can only open ten databases a time, but this will be upgraded to open as many as needed. Only two DBMSs are supported, but other widely-used ones will be supported. The tool cannot import from other bibliographic sources. The main dialog cannot be resized. A Web-based application is not provided in the current version. The current status of the development is alpha, so that several bugs need to be fixed, as well as some non-intuitive behavior. In addition to this short list, a given user might find several other limitations. We have only highlighted the most noticeable ones (from our point of view).

## 2.2 Technology

Bib Manager has been completely implemented using Java under Eclipse. MikTeX 2.5 [8] has been used for generating reference keys along with OS batch files.

## 2.3 Description of Bib Manager

Figure 1 shows the main GUI of Bib Manager, which consists of a row/column-based table view for several databases, organized in a tab form (left-hand panel). The right-hand panel allows filtering of entries that contain the typed text in any of the fields Author, Title, Year, and Key. The same is also possible for the rest of the fields in the text box Other Fields. Note that the filtered entries contain the typed text somewhere in the required field, which is useful, for instance, if we are looking for the publications of an author who may also be a co-writer. Also, in this right-hand panel, the ordering of entries can be selected for any field, which is also useful for lookups. The number of displayed references (which depends on the applied filter) is counted in the text box Ref's. The bottom panel includes icons as shortcuts for frequent operations; from left to right: the magnifying glass for switching BibTeX and table views (see Figures 1 and 2), the trash can for deleting selected entries, the opened book for inserting a new reference, the selection of a book for modifying a single entry, the tools for configuring preferences, and the double arrow for copying selected entries between two databases. Since data read from the database are kept in main memory, updates from other users are not seen unless a view refresh is done (the button Refresh View is missing but

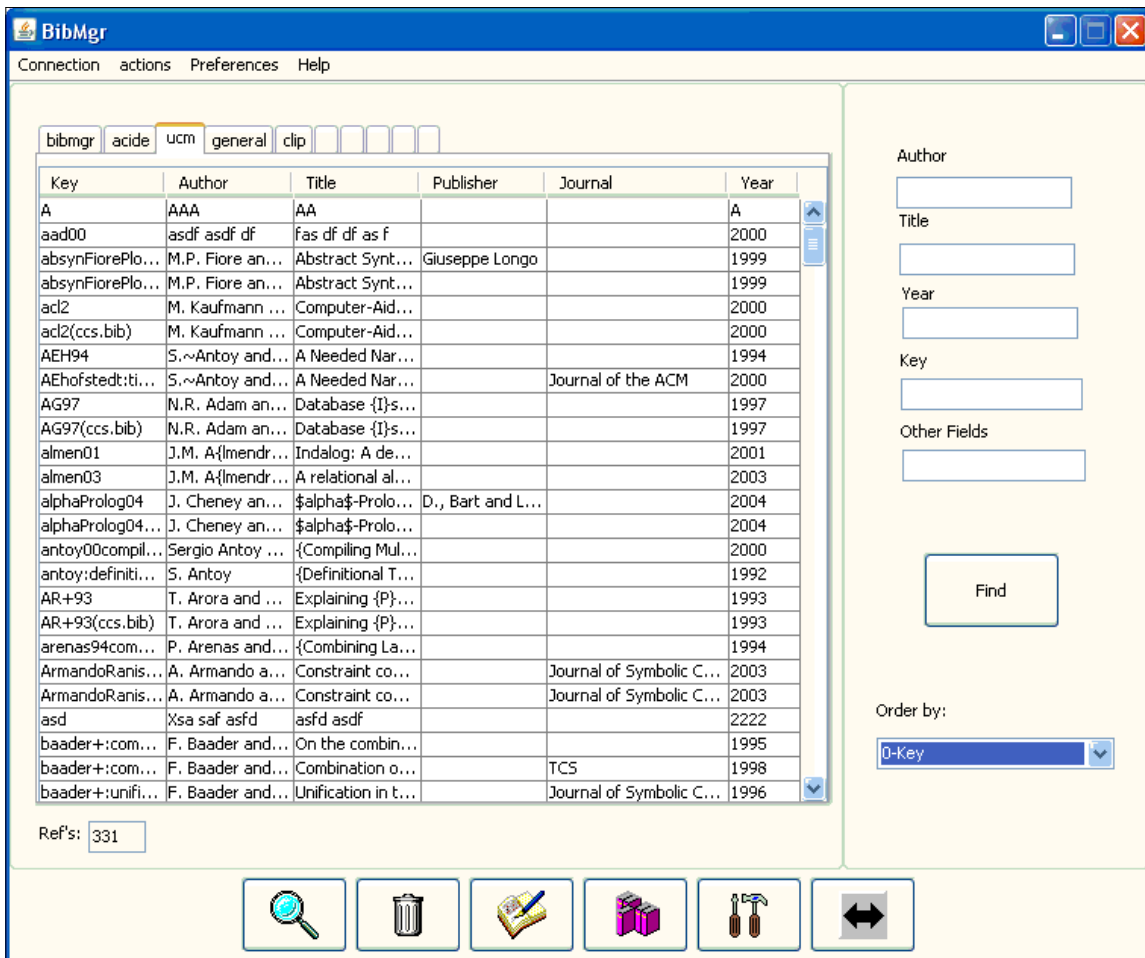


Figure 1: Bib Manager Main GUI

under development). In any case, updates from other users are supervised by the DBMS, therefore avoiding duplicates because the primary key of the database table is the B<sub>B</sub>T<sub>E</sub>X key.

The menu bar has several entries:

- Connection. For connecting to and disconnecting from both MySQL and Access databases. Connection names as used in Word Citer (see Section 3.3) should be added.

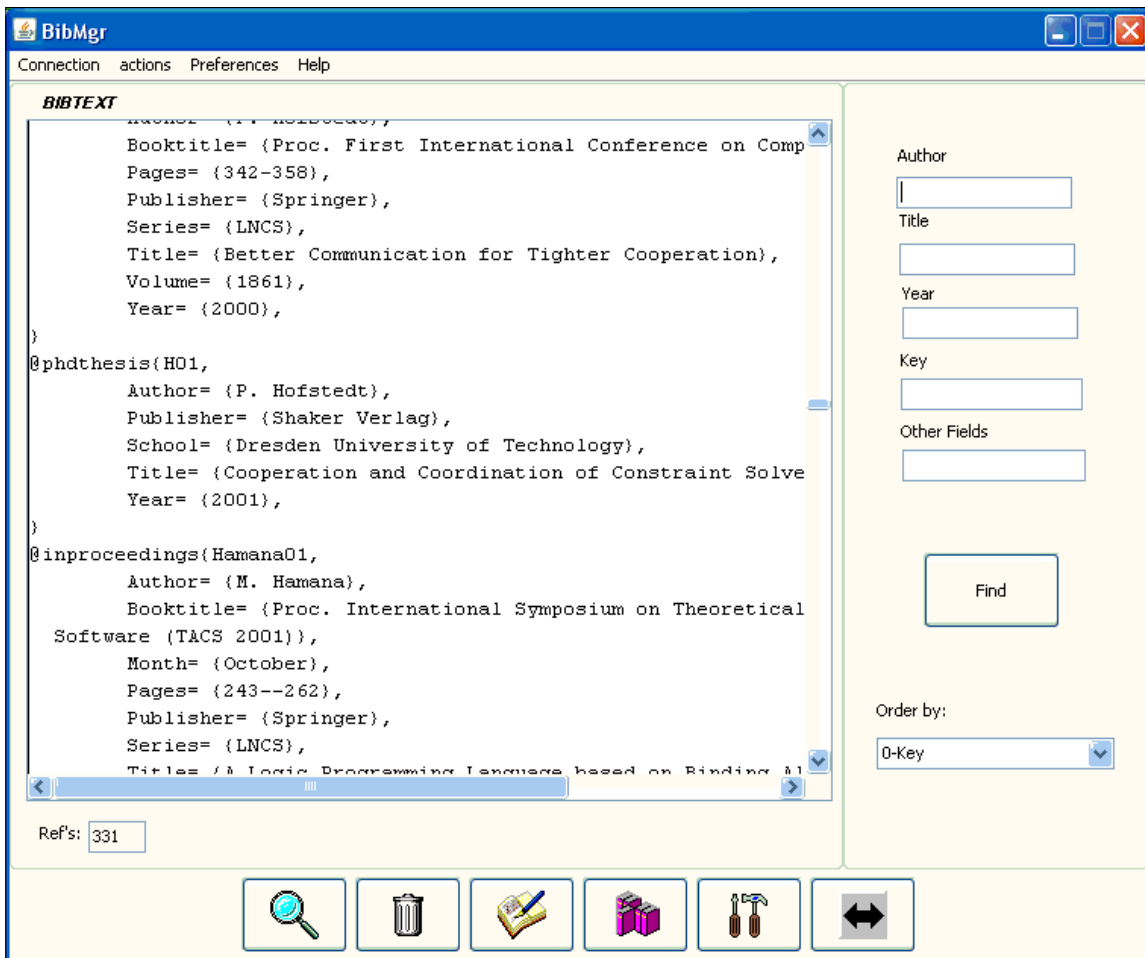


Figure 2: BibTEX View in Main GUI

– Actions:

- Insert Reference. Opens a dialog box for inserting a new reference (see Figure 3). Depending on the entry type (book, article, inProceedings, etc.) that had been selected in the drop-down box, the dialog will show the corresponding mandatory and optional fields to be filled. It is possible to automatically generate a key from the citation label that L<sup>A</sup>T<sub>E</sub>X/BibTEX generate in terms of the filled fields (currently, only the style alpha is supported). For this, both a .tex document containing

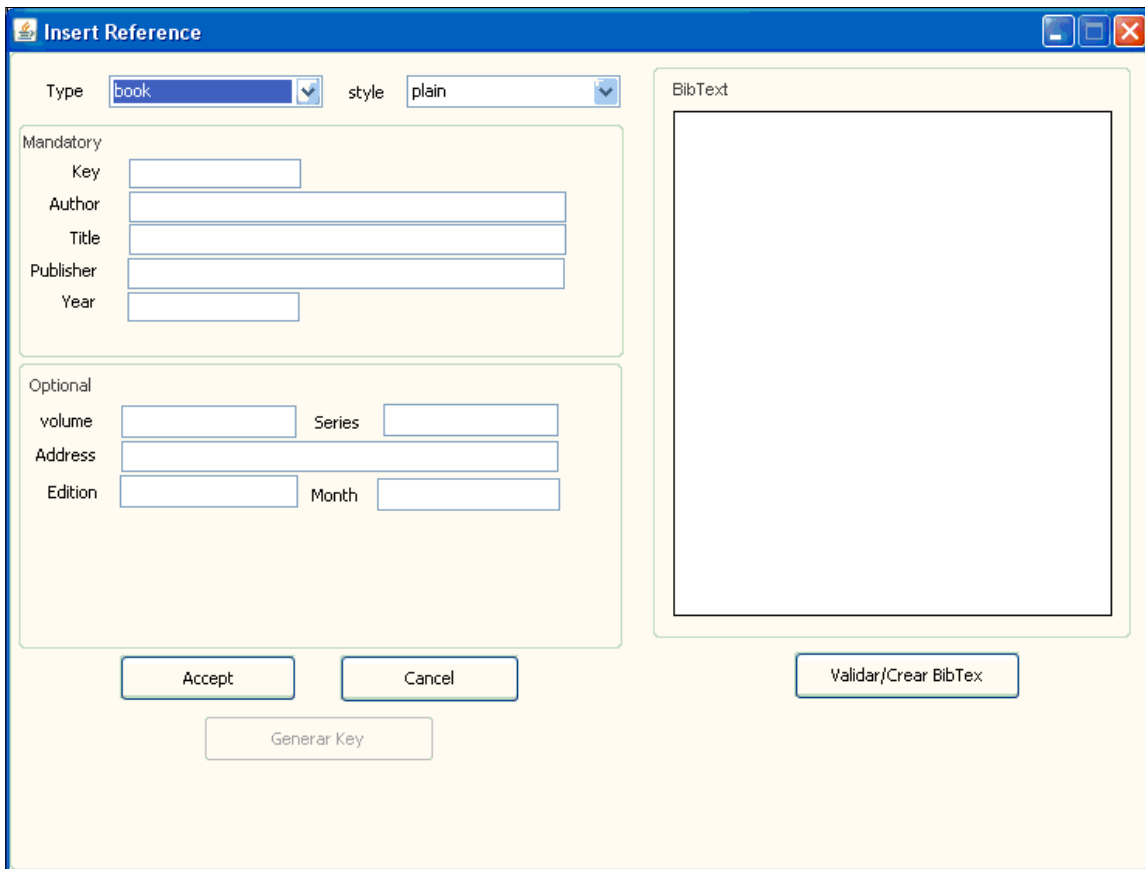


Figure 3: Inserting a Reference

a `\cite{gen}` and a `.bib` containing a `BIBTEX` entry are generated and compiled. The result of the compilation contains the automatically generated citation label, which is extracted from the `.bb1` file.

The button Create/Validate Entry allows two operations. First, to display in the text box Bib Text the `BIBTEX` text format corresponding to the filled fields of the reference (left-hand panel). Second, to automatically fill the fields of the reference from a manually typed entry (or pasted from another source) in the text box Bib Text (this last point is under development).

- Delete Reference. Deletes the selected references. The key Del is a short-



cut.

- Copy Reference. Copies selected references to another opened database, which can be from any supported vendor. Ctrl+C and Ctrl+V should be shortcuts, which is a future enhancement.
- Modify Reference. Opens a dialog box for modifying an existing reference (see Figure 4). The dialog is the same as the one for inserting

The screenshot shows the 'Modify Reference' dialog box. It features a title bar with the text 'Modify Reference' and standard window controls (minimize, maximize, close). The dialog is organized into several sections:

- Type and Style:** Two dropdown menus at the top. 'Type' is set to 'article' and 'style' is set to 'alpha'.
- Mandatory Fields:** A section containing five input fields:
  - Key:** AEhofstedt:tigher-coc
  - Author:** S.~Antoy and R.~Echahed and M.Hanus
  - Title:** A Needed Narrowing Strategy
  - Journal:** Journal of the ACM
  - Year:** 2000
- Optional Fields:** A section containing four input fields:
  - Volumen:** 74(4)
  - Number:** (empty)
  - Pages:** 776--822
  - Month:** (empty)
- BibText:** A large text area on the right containing a BibTeX entry:

```
@article{AEhofstedt:tigher-cooperat
  key_ = {AEhofstedt:tigher-coopera
  title= {A Needed Narrowing Strat
  author= {S.~Antoy and R.~Echahed
  journal= {Journal of the ACM},
  year= {2000},
  volume= {74(4)},
  pages= {776--822},
}
```
- Buttons:** Four buttons are located at the bottom: 'Accept', 'Cancel', 'Validar/Crear BibTex', and 'Generar Key'.

Figure 4: Modifying a Reference

a new reference, but here we can see other fields because a different reference type is selected. Also, double-clicking on the table view edits the contents of a cell.

- Create Database. Creates a MySQL or Access database.

- Delete Database. Deletes (drops) a MySQL or Access database. The key Del should be a shortcut when the focus is in a tab.
- Import BibTeX. Imports BIBTEX entries from a .bib file.
- Export to BibTeX. Exports the selected database to a new .bib file.
- Preferences. Sets the user preferences for hidden/displayed columns, the GUI language, and the paths of the executables for L<sup>A</sup>T<sub>E</sub>X and BIBTEX.
- Help. Contains BibMgr Help, Log to show the log, and About Bib Manager.

This completes a short introduction to Bib Manager. The next section will describe a citer for MS Word documents, Word Citer, which can use the databases managed by Bib Manager.

## 3 Word Citer

### 3.1 Features and Limitations

Word Citer is a tool that resembles the synergy between L<sup>A</sup>T<sub>E</sub>X and BIBTEX, with added capabilities. It allows inserting of references from different sources: Bib Manager databases and BIBTEX files. There is provision for useful searching of references, similar to the search described in Section 2.3. When a new reference is first inserted in the Word document, its key is used as the unique identifier. Later, when listing the complete set of references, citation labels are resolved and the citations are updated with these. Citation labels appear following the desired bibliography style, whether predefined or custom. In this last case, any .bst valid file can be selected. In addition, the format of the references in the listing follows this bibliographic style.

There are several limitations in this tool. First, only MySQL databases can be handled, which has to be enhanced to support also MS Access, and several others in common use (PostgreSQL, Oracle, Sybase, DB2, etc.). Second, it cannot handle other bibliographic sources such as the ones that can handle EndNote [5], ProCite [10] or BibShare [3]. Similar to Bib Manager, there are several other limitations which we are tracking and hope to improve.

### 3.2 Technology

Word Citer has been completely implemented using Visual Basic for MS Word 2003. MikTeX 2.5 [8] has been used for generating the list of references and citation labels. OS batch files have been also used.

### 3.3 Description of Word Citer

Figure 5 shows the new entry in the Word menu bar after the installation of Word

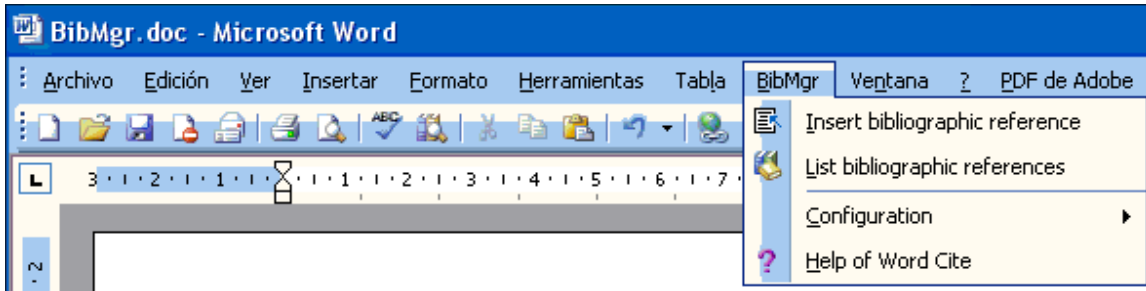


Figure 5: Menu Bar

Citer. The third menu item of BibMgr allows configuration of Word Citer (see Figure 6). The very first task a user will do with Word Citer is to configure the data

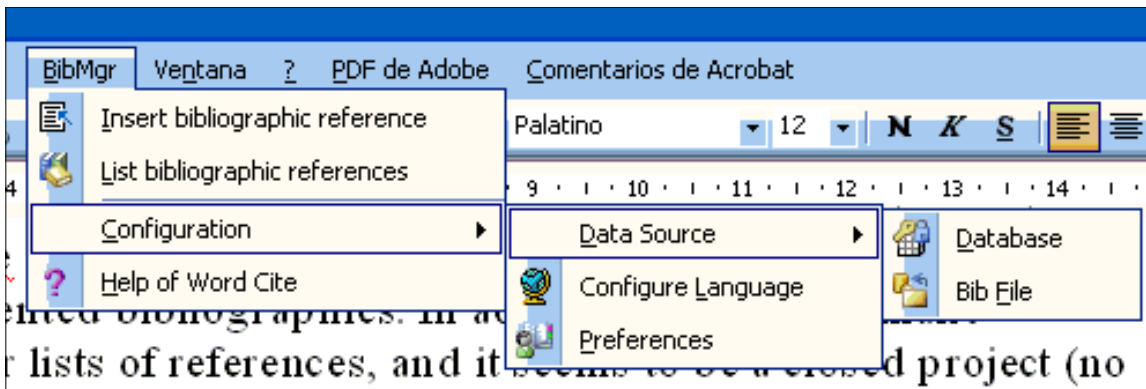


Figure 6: Menu Item for Configuring Word Citer

source, which can be a MySQL database or a .bib file, as shown in the figure. If the user selects the latter, an open-file dialog box appears for selecting the requested file. This operation will create an in-memory dummy database, hidden to the user, and handled as any other database regarding reference lookups and insertions. Selecting such a .bib file allows the user not only to use exclusively the references in this file, but also to add the new references to the existing ones, making it possible to handle several .bib files in the same Word document (as it is usual in L<sup>A</sup>T<sub>E</sub>X documents). If the user otherwise selects a database as a data source, the dialog box in Figure 7 appears. With this dialog, the user can manage

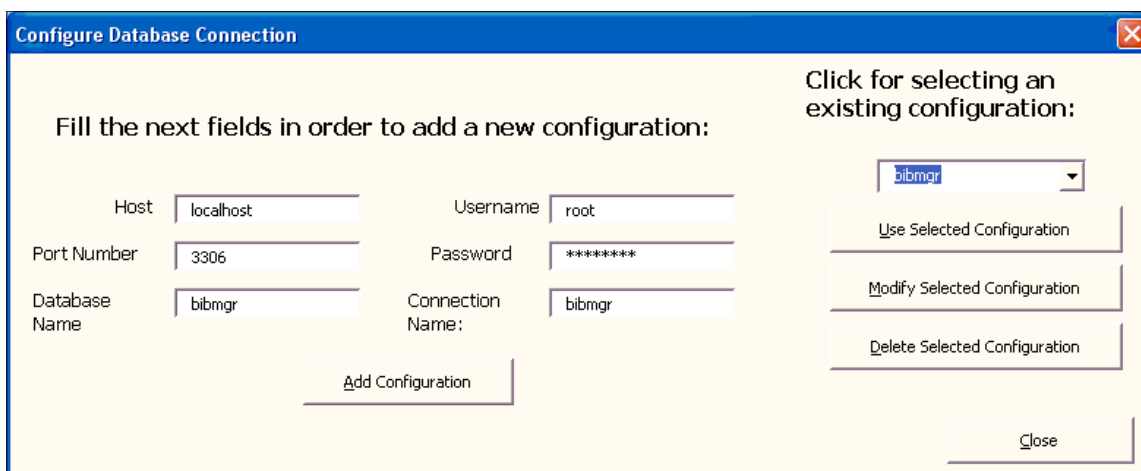


Figure 7: Selecting a MySQL Database Source

connections, which are identified with a connection name. The user can select any of the previous connections or create a new one. Whenever Word is started, the last connection is used, therefore saving time when continuing previous work. Connections can be managed with usual operations: create (Add Configuration), delete (Delete Selected Configuration), and modify (Modify Selected Configuration). To connect to a saved connection, the user pushes the button Use Selected Configuration. The list of previous connections is stored for future use in the drop-down box next to the upper-right corner. A straightforward enhancement is to allow the use of several database sources to look for references at the same time.

The second task the user will usually do is to insert a new bibliographic reference using the menu item Insert Bibliographic Reference, which opens the dialog

shown in Figure 8. In this dialog, the complete set of references from the chosen

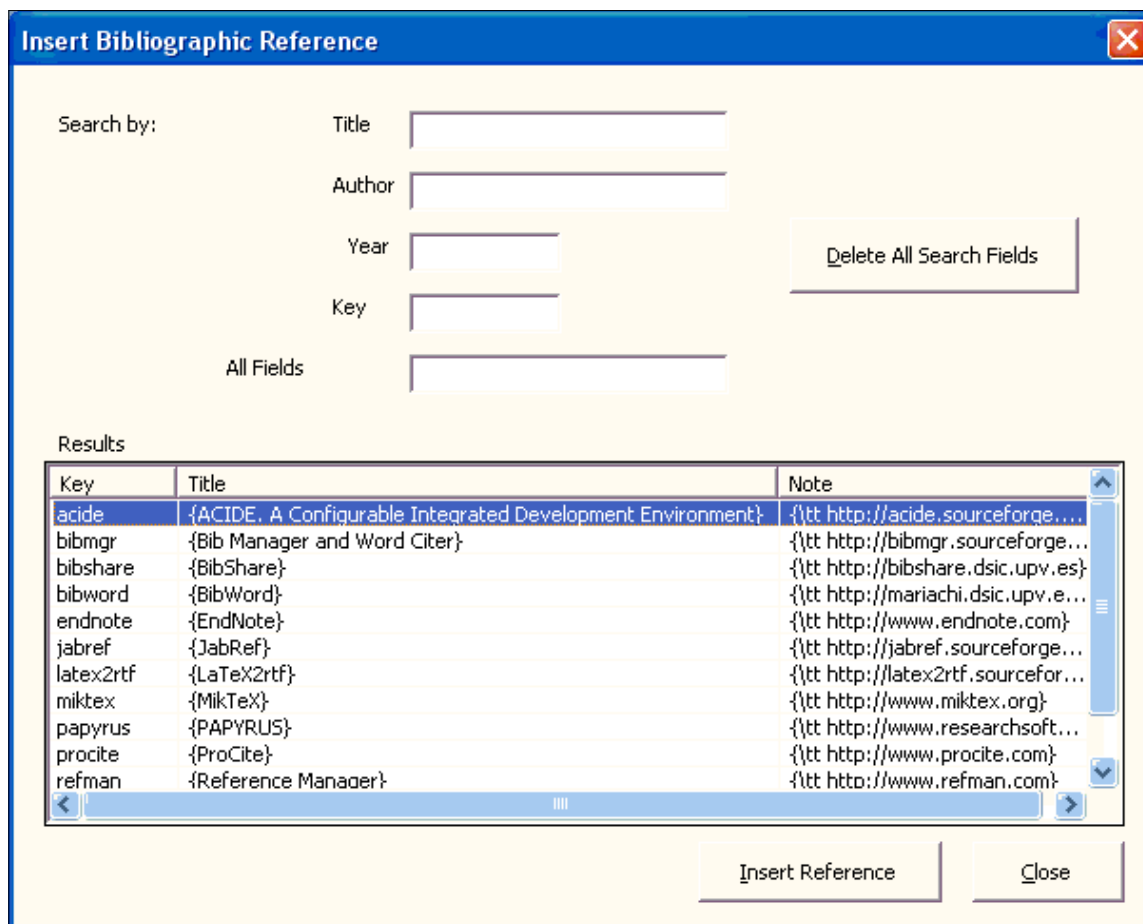


Figure 8: Inserting a New Reference

data source is shown in a table view. Filtering can be done similar to the method explained in Section 2.3 for facilitating the selection of the required reference. For instance, if we type bib in the Title search field, we get the result shown in Figure 9. Recall that the search text is looked for in any position of the database field.

Inserting a concrete reference amounts to inserting a new Word field at the current cursor position, which displays the  $\text{BIBTEX}$  reference. When the list of references is created afterwards, this Word field will display the actual citation label generated by  $\text{LATEX}$  and  $\text{BIBTEX}$ . Figure 10 shows the result of inserting a

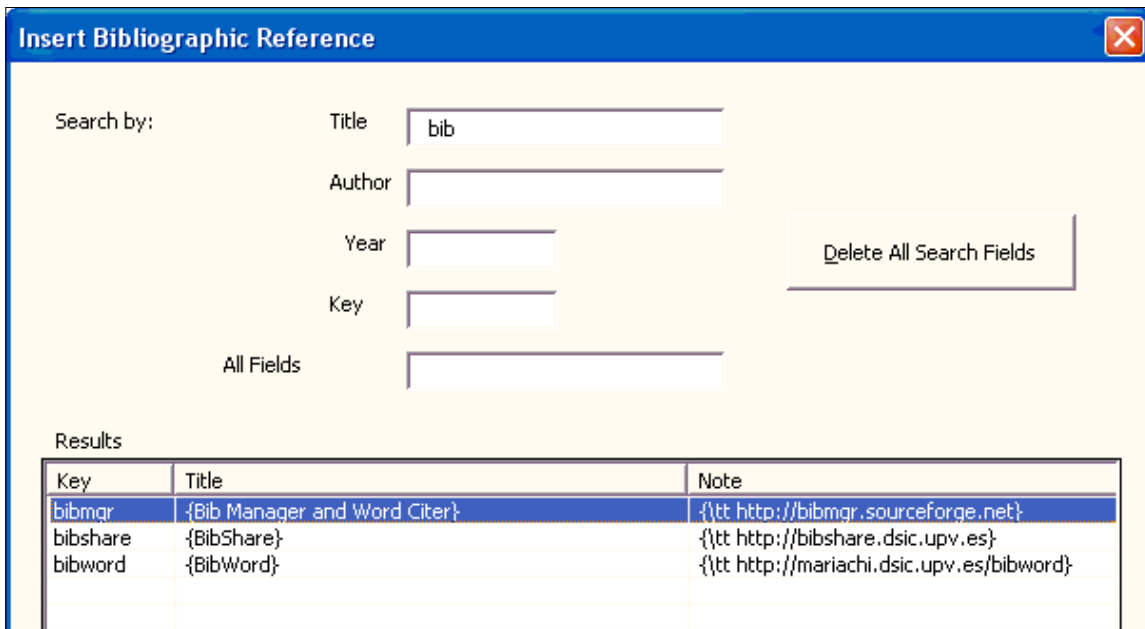


Figure 9: Filtering References

reference.

The last task is to list the references in the document, which is done with the menu item List Bibliographic References. Figure 11 shows the result, which is based on the `abbrv` bibliography style. Currently, this operation can be done only once, a known bug which will be soon fixed. `LaTeX2RTF` [7] is used to generate an RTF document which is the source for the listing.

Several configurations are available for Word Citer. As seen in Figure 6, in addition to configuring the data source, the GUI language can be selected from a drop-down list (resource files are used to define all the texts in the dialogs) as well as some preferences (Preferences, see Figure 12). These refer to the columns that the user wants to display when inserting references in the table view, the paths for `LATEX` and `BIBTEX` executables, and the bibliography style (a built-in or any valid `.bst` file).

This concise introduction to Word Citer closes this section. In the next section some conclusions and future tasks are outlined.

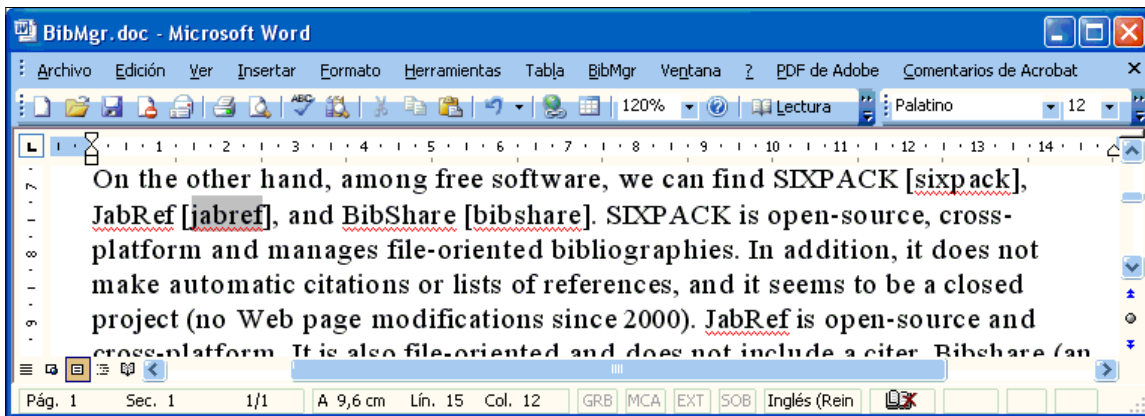


Figure 10: The Result of Inserting a Reference

## 4 Conclusions and Future Work

This paper has presented a bibliographic manager and citer, which are alternatives to other available tools, with features meeting our initial requirements. They cannot be thought of as complete tools since they are emerging and in an alpha development status. They should be seen as tools that, in time, might provide more and more features following our requirements. There are other tools in the free market today, but Bib Manager and Word Citer might be competitive in future development stages.

There are many goals to be achieved, some of which have been posed in the sections describing the tools, and others which have been indicated or have become evident after each tool description. Apart from fixing numerous bugs (including operating system dependent calls, which makes Bib Manager work only for Windows, up to now), we have to implement other interesting features such as the handling of other databases and even on-line access to bibliographies on the net, according to standards. Despite the assumed limitations, we hope that both Bib Manager and Word Citer tools will consolidate into friendly and powerful applications, amenable to its targeted users; in particular, to L<sup>A</sup>T<sub>E</sub>X/T<sub>E</sub>X and/or MS Word users. This project may be downloaded from <http://bibmgr.sourceforge.net>. We hope to upload the version described in this paper as soon as possible, and have a better and stable implementation.

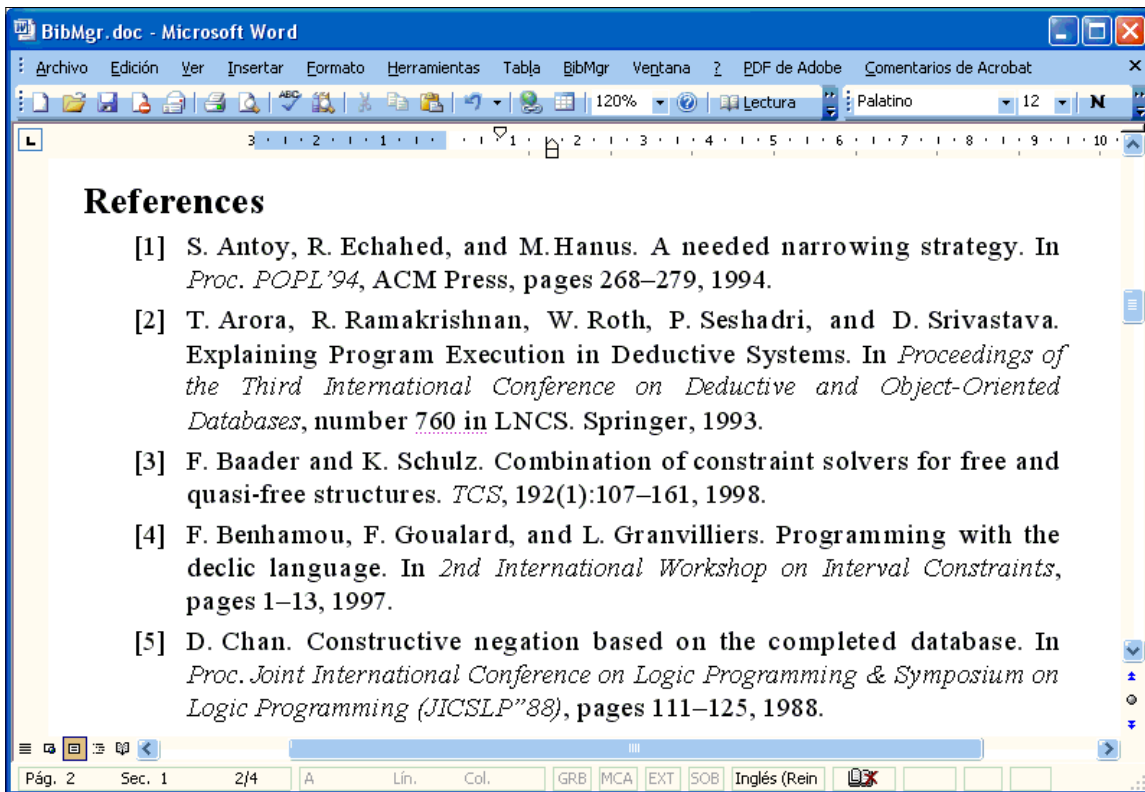


Figure 11: Reference Listing

## 5 Acknowledgements

I am grateful to the students who developed this system under my guidance, which I hope will be useful for them in the development of their own future projects. These students are Antonio Contreras Varas, Enrique Palomar Santamaría, and Javier Prats Robledillo, who developed this project in the computer science course “Sistemas Informáticos” (Computing Systems) for graduates, a course intended for the fifth year of Computer Engineering studies of the Computer Science Faculty of the Universidad Complutense de Madrid, Spain. Also, thanks to the projects TIN2005-09207-C03-03, and S-0505/TIC0407 which supported this work. Finally, thanks to the free, open-source project LaTeX2RTF [7] which we use in our project, as well as to the Java and Eclipse initiatives. This



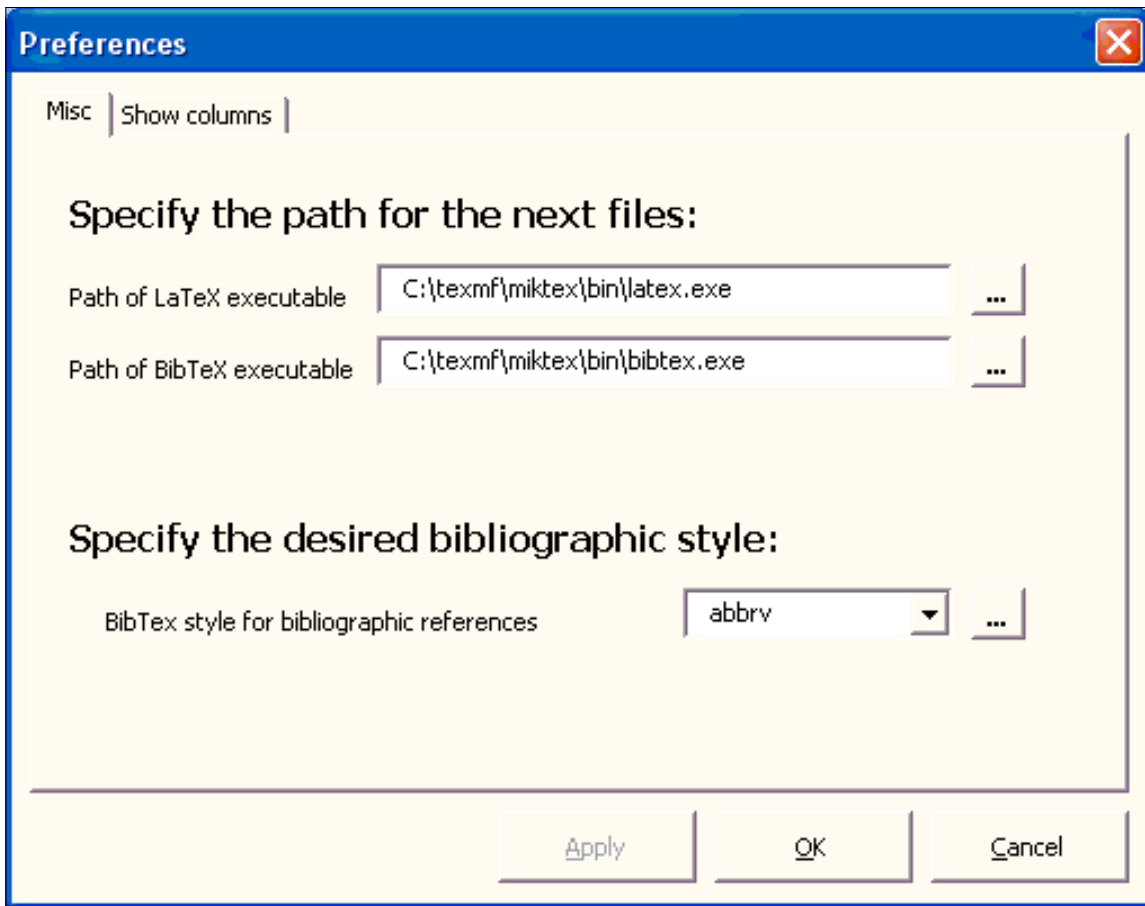


Figure 12: Word Citer Preferences

article was prepared with MikTeX 2.4 [8], ACIDE [1], and BibMgr [2].

## References

- [1] ACIDE. A Configurable Integrated Development Environment. <http://acide.sourceforge.net>.
- [2] Bib Manager and Word Citer. <http://bibmgr.sourceforge.net>.
- [3] BibShare. <http://bibshare.dsic.upv.es>.

- [4] BibWord. <http://mariachi.dsic.upv.es/bibword>.
- [5] EndNote. <http://www.endnote.com>.
- [6] JabRef. <http://jabref.sourceforge.net>.
- [7] LaTeX2RTF. <http://latex2rtf.sourceforge.net>.
- [8] MikTeX. <http://www.miktex.org>.
- [9] POPYRUS. <http://www.researchsoftwaredesign.com>.
- [10] ProCite. <http://www.procite.com>.
- [11] Reference Manager. <http://www.refman.com>.
- [12] RefWorks. <http://www.refworks.com>.
- [13] SIXPACK. <http://sixpack.sourceforge.net>.