4

USING BIBLIOGRAPHIC DATABASES

Dr. Meredith: A bit of advice …
Mitch: Oh, uh, thank you …
Dr. Meredith: Always … no, no … never … forget to check your references.
—Real Genius (1985)

4.1 IN THIS CHAPTER

This chapter explains the general issues of using a bibliographic database, including how team members can share and contribute to a bibliographic database as they collaboratively manage the citations and resources associated with a project.

A big difference between references and all the other nontextual elements of a document (mainly headings, figures, and tables) is that managing references requires a steeper learning curve. With other elements, you can download a *STREAM Tools* template, and after a brief instruction on how to "reuse the templates," you are ready to go. Incorporating references, however, requires a separate software package and substantial learning time. If the number of references you are dealing with is small, forget about the specialized software and process them manually. For example, if you are a plant floor

Technical Writing for Teams: The STREAM Tools Handbook, by Alexander Mamishev and Sean Williams

manager and you need to reference a couple of manuals and a newspaper article, just type them in using plain text or footnotes. If, on the other hand, you are a member of a research organization and you need to write a journal paper, a Ph.D. thesis, a large federal agency proposal, or a book, then the time you invest in learning reference management techniques will be returned to you and your colleagues many times over.

4.2 WHY USE A BIBLIOGRAPHIC DATABASE?

There are several reasons for using a bibliographic database.

First is the need for automatic numbering of your literature citations and automatic ordering of the citations at the end of the document. Many journals require that your literature citations be numbered consecutively in the order they appear in the document. In this case, the text would look like this: *"Although bird worms were proven to be affected by zero gravity [22], it has been demonstrated that alien forms are not susceptible to these conditions [23]."*

Second is the need for automatic formatting of the citations list. The expected format is different for each journal. For example, some journals require that the citations appear alphabetically and are referenced by the name of the first author and the year they were published. The previous example would then look like: *"Although bird worms were proven to be affected by zero gravity [Dinkley2003], it has been demonstrated that alien forms are not susceptible to these conditions [Lawrence2004]."*

Third is the opportunity to keep your references organized by your own categories. You may want to be able to generate a quick list of all papers published in your research group in the last five years, or a list of all papers that make new contributions to the body of knowledge on the environmental impact of a certain technology, or a list of all papers by a competing group. Some of these lists will be an ongoing effort, to which several people on your team should contribute. For example, if you want your group to write a review article on research in your field, you could ask every group member to add or mark papers in the database that they consider important in the course of their daily activities in the next six months.

Fourth is the ability to share references. For example, say a graduate student just finished a dissertation, which comprises a 250-page manuscript with 350 references. The next graduate student is continuing the work, starting from scratch. Standing on the shoulders of giants does not mean doing things from scratch.

Managing hundreds or even thousands of references manually in light of all the above issues is an extremely daunting task. It quickly turns into busy work of no practical value. The following section presents additional justifications for using bibliographic databases and the basics of using the most common tool for citation management: EndNote.

4.3 CHOICE OF SOFTWARE

If you adopted Microsoft Word as your flagship writing software, you could, in principle, use the built-in capabilities of the **Citations & Bibliography** ribbon in Microsoft

Word 2007. However, the built-in capabilities of this option are so limited that we cannot recommend it for collaborative writing, partly because authors cannot share these resources among team members since the database of sources resides on a single computer.

Three main database packages exist that are appropriate for collaboration: EndNote, Reference Manager, and ProCite. EndNote is the most popular and easiest to use, and we recommend it as your first choice. Reference Manager comes in at a close second; it has sophisticated office network features, allowing multiple users to access multiple databases at the same time. The problem is that Reference Manager isn't compatible with individual EndNote users, who could be scattered everywhere and do not necessarily belong to your organization. Importing or exporting between EndNote and Reference Manager is not overly difficult, but it is an annoying extra step, so be sure to select your initial software carefully. If you have a powerful IT department and highly streamlined computer support, the network advantages of Reference Manager might outweigh the fact that it is less popular than EndNote. The third option, ProCite, is a close relative to the other two. It does not offer clear advantages over either, however.

If you need to interface to LaTeX, you will also need BiBTeX. EndNote, Reference Manager, and ProCite are bibliography management systems to be used with Microsoft Word, and BiBTeX is the bibliography management system for LaTeX. Since the differences between the first three are almost cosmetic, you may want to make your choice based on what your potential collaborators are already using. After that, you may want to operate them in such a way that you maintain compatibility with LaTeX.

In short, if you are not sure what to pick, use EndNote. You will need to choose between the web-based and the desktop-based modes. The following instructions correspond to the desktop version.

4.4 USING ENDNOTE

This section shows an example of using EndNote database in a research group.

4.4.1 Setting Up the Interface

Install the EndNote software. The website www.endnote.com offers a 30-day software trial, as well as multiple tutorials and webinars teaching the more advanced functions of EndNote. The following example will cover the basics of getting started, as well as a discussion of the applications available for using EndNote in teams. You may need to supplement this info with additional instructions and tutorials selected based on your specific needs.

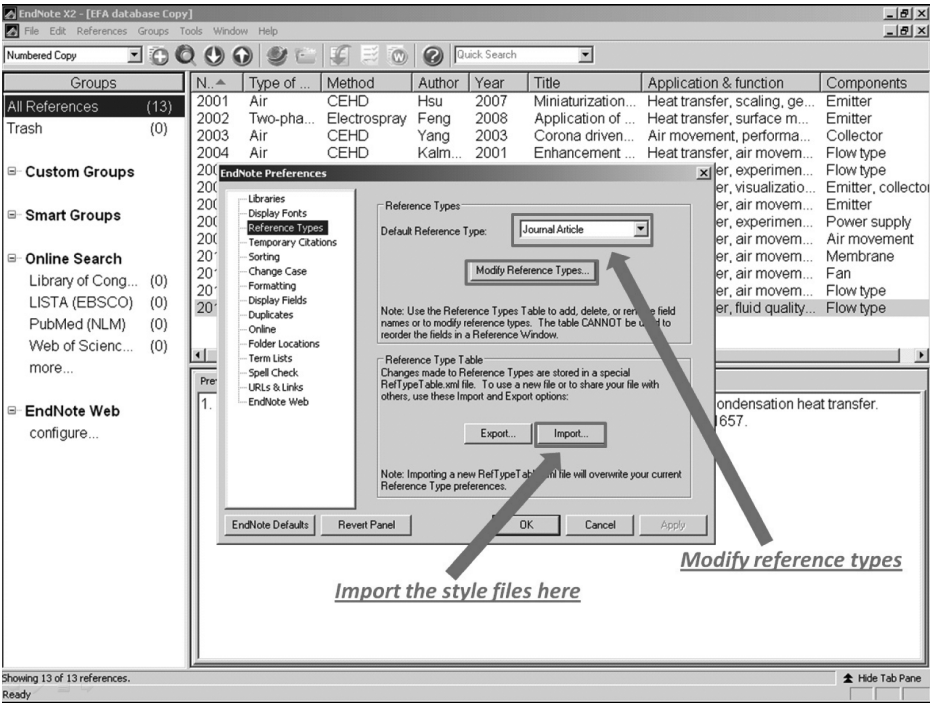
Once your software is installed, you can either select an existing database (perhaps created by your colleagues) or create a new one.

To open an existing database, click **File** → **Open** → **Open Library**, and then select an existing database.

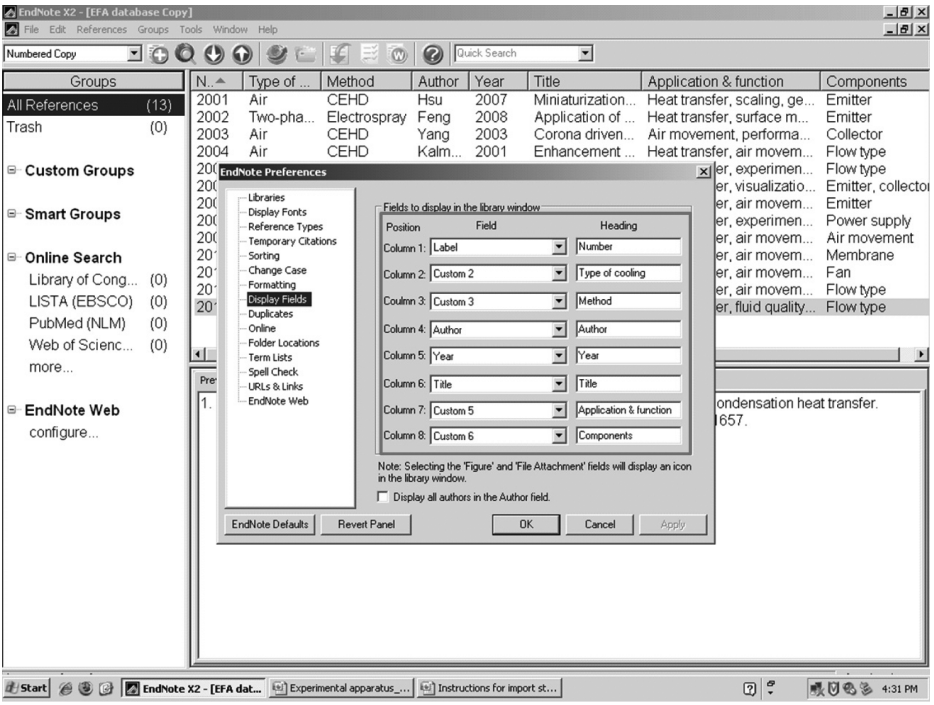
To create a new database, click **File** → **New File**. Note that when you create a new file, EndNote will generate a **.enl* file and **.Data* file folder. These two, the file and the file folder, should be kept together in order to open attached files from references.

Your next step is to customize the database interface. Specifically, you need to decide which database field entries need to be displayed, which can be omitted, and which custom fields need to be created to meet your needs. Frequently used custom fields are Author, Title, Year, etc. Some groups may want to include classification terms for each paper in their database; some need a custom field for conversion to BiBTeX; some may want to link each entry to a full text .pdf file, and so on. Each user and each group, of course, will have their own preferences for the interface, which will evolve over time and with more collaborative use. To illustrate the importance and the convenience of interface customization, we provide an example of a customized database that can be shared among multiple users.

To see the example, visit the website streamtoolsonline.com and download the file titled "StandardInterface." This is a reference style file; it controls the appearance of the interface in EndNote. In order to add an existing reference style, click **Edit** → **Preferences**, select "Reference Types," and then select the desired file (in our case, "StandardInterface") from the directory. See the screenshot below for this step.

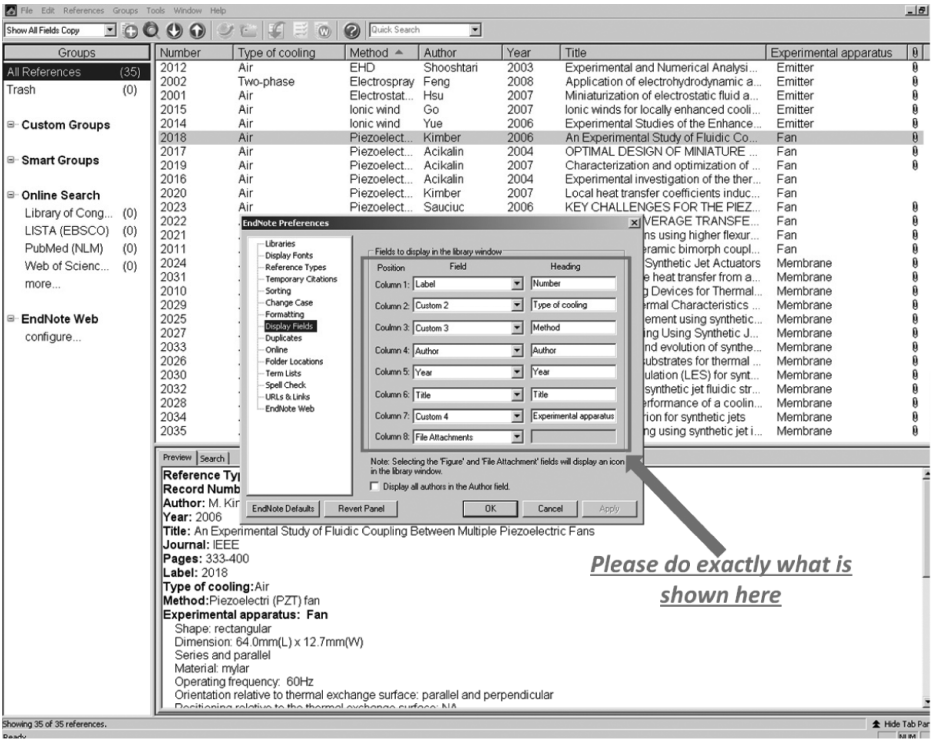


Your next step is to mark "Display Fields" and select fields that you want to display. The screenshot below and the following table show a practical selection of custom-defined columns for a mid-size project.



Position	Field	Heading
Column 1	Label	Number
Column 2	Custom 2	Device type
Column 3	Custom 3	Method
Column 4	Author	Author
Column 5	Year	Year
Column 6	Title	Title
Column 7	Custom 5	Application & function
Column 8	Custom 6	Components

Having defined the characteristics that are important to your group, you may choose to develop a detailed taxonomy for your field of work.



4.4.2 Adding References

Now you are ready to add references. There are several sources for your references:

a. You can find them online from inside the EndNote software. To achieve that, click **Tools**, **Online Search**, then select the database in your field from the list. This will take you to a classic search interface, where you can search by author's name, words in the title, etc. After completing the search, you can select references of interest to you, and move them to your permanent database.

b. You can find them online by using your web browser with access to online databases, or, for example, Google Scholar and then enter the information manually.

c. You may be able to find references in library databases and then export them to your EndNote software with a few mouse clicks if the database allows this type of function.

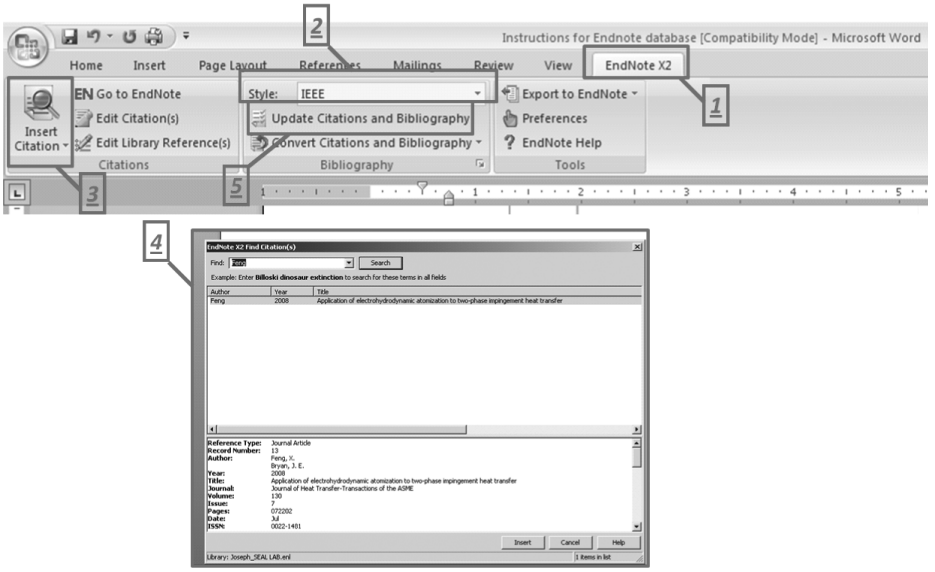
d. You may obtain reference collections from your collaborators (and if they use Reference Manager, you can import their files into your EndNote database).

e. You can acquire references previously entered by your group members who are working on the same project. A more detailed discussion of this method comes later.

4.4.3 Citing References

Once you have created the database, citing references is easy. If you use Microsoft Word 2007 and have already installed EndNote, you will see a dedicated EndNote ribbon on your screen, as shown in the screenshot below. After selecting the EndNote ribbon, select the Style for your references (e.g., *IEEE* requires one style, *Nature* requires a different style). It should be noted that, unfortunately, many of the provided styles are placeholders rather than real styles. In other words, their output is quite different from the actual publisher requirements and much work is needed to modify the style files to produce the officially required output. We provide some of the corrected styles at streamtoolsonline.com, and hope that perhaps with time the software makers will update the style files that they provide.

Click **Insert Citation** → **Find Citation**, after which you will be switched to EndNote. Then search and select the citations you want, and click insert. As the last step, click **Update Citations and Bibliography**.



These instructions provide you with a birds-eye view of main functions that you need; we also encourage you to review the free online tutorials available at endnote.com, partly because of the interactive multimedia delivery mode and partly because newer versions of the software may change some of the functions. It is not likely, however, that any major changes will be made, because EndNote is a mature software package.

4.5 SHARING A DATABASE

A research group that works closely in the same field may want to share the same database, so that the references entered by one person can be reused by others in the future. In addition, the group member who entered that reference may choose to add comments, something like "*this is a very important paper for us because it describes the methodology for the numerical modeling of thermal diffusion*." An even more sophisticated approach is to create a classification scheme for relevant papers.

4.5.1 Numbering the Database Entries

To achieve citation numbering, assign each reference in the database a number. If several people want to use the same database in their work, they should be able to enter new references at the same time. To achieve that, the research division is broken into groups or individuals and each group or individual is assigned a block of numbers in the shared reference database. For example, John has the block 201–300, Julie has 301–400, and Jill has 401–500. At the beginning of the project, each user downloads the master database to their computer. Now they have a clone database, which has all entries from earlier dates and the new entries entered by the user in his or her block. Since everyone will enter new references in their block, they will not collide when the database is synchronized later. To synchronize the database, the user downloads the most recent master database to their computer, copies the new entries from the clone to the master, and uploads the new master database to the central location. As an alternative to the periodic synchronization approach, everyone could use the same file that resides on a shared drive, if such a capability is provided by the organization.

There are no special tools in EndNote to assign blocks of numbers for use. This is accomplished simply by maintaining a list at a common location. After all the numbers in the block have been used up, a new block is assigned.

In our experience, even if the system is implemented in the group, many participants choose to ignore these assigned blocks of numbers, starting their database from number 1 because, they believe, it is just easier that way. These participants pay the price later, when they have to synchronize with the master database under the peer pressure of the more efficient group members who have worked within the assigned numbering scheme.

A few simple rules to ask of your group participants:

1. Stay in your block of numbers, even if it means that you have to assign the number to each new reference manually.
2. In order to modify entries outside of your block (e.g., because you found errors) you need to check out the master database—otherwise your changes will be lost during database synchronization.
3. Make sure to check for duplicates—it is likely that someone already found your entry. On the other hand, duplicates do not present any problems, and if a few slip by you, do not worry about it.

4. It often makes sense to put papers generated by the group members into a dedicated block. As team members publish new papers, they add them to this block. This will make it easier to reference prior work.

4.5.2 Compatibility with BiBTeX

Early on, you will want to determine whether you need to maintain compatibility with LaTeX. The software module in LaTeX that manages literature references is called BiBTeX. In BiBTeX, each reference has a field for unique identification. The most common way to identify a paper in the BiBTeX database is to use the first author's name and the year of publication. Once you dedicate one of the custom-defined columns in the database for this field, you will be able to conduct import-export actions between BiBTeX and EndNote or Reference Manager.

The procedures for adding custom fields in EndNote are described above. To display user-defined columns, right-click the **RefID** tablet and select **Reference List Display**. Increase the number of columns viewed and select the columns that you want to be visible.

For example, consider the following paper:

C. P. Hsu, N. E. Jewell-Larsen, I. A. Krichtafovitch, S. W. Montgomery, J. T. Dibene II, and A. V. Mamishev, "Miniaturization of Electrostatic Fluid Accelerators," *Journal of Microelectromechanical Systems*, vol. 16, no. 4, pp. 809–815, Aug. 2007.

The entry in the column "User Def 1" would be *Hsu07*, and if you expect to see more papers for the same author, you can add the first word of the title, so that the entry would look like this: *Hsu07Miniaturization*.

Another facet of compatibility with others is providing readily formatted database entries to others. Some academic authors choose to put readily-formatted BiBTeX and EndNote collections of references to their own work on their websites, thus inviting web page visitors to cite their work. While this practice could be considered slightly aggressive and self-serving, it is entirely within the ethical bounds of the academic community.

4.6 FORMATTING REFERENCES

Maintaining the proper formatting of references according to publisher requirements is a daunting task, as there are literally thousands of formats. Although the style files discussed in Section 4.4.1 reduce the total amount of formatting efforts and the output quality, inexperienced writers tend to make a large number of formatting mistakes when they cite literature, simply because they are completely unaware of the existing conventions. Some of these conventions are quite universal, and some apply only to narrow fields of specialization. Since the number of formatting rules is so large and they vary so much, a writer should develop a certain intuition about formatting requirements.

The following examples intend to provide initial training on this subject for inexperienced writers.

Find typesetting and stylistic mistakes in the hypothetical text below:

A comprehensive overview of interdigital sensors and transducers is provided in [1].

…

One of the earliest examples of using patrolling robots for sensing properties of electric power cables (2) demonstrated technical feasibility of autonomous mobile sensing for maintenance of distributed infrastructures.

…

Computationally intensive algorithms that provide a high rate of automated detection and discrimination of a broad range power quality events [e.g. those described in [3,4] can now be implemented in hand-held diagnostic devices.

References:

[1] A.. V. Mamishev, K. Sundara-Rajan, F. Yang, Y. Q. Du, and M. Zahn, "Interdigital Sensors and Transducers," *Proceedings of the IEEE,* vol. 92, no. 5, pp. 808-845, May 2004.

[2] C. P. Hsu, N. E. Jewell-Larsen, I. A. Krichtafovitch, S. W. Montgomery, J. T. Dibene II, and A. V. Mamishev, "Miniaturization of Electrostatic Fluid Accelerators," *Journal of Microelectromechanical Systems*, vol. 16, no. 4, pp. 809-815, Aug. 2007.

[3] M. Wang and A. V. Mamishev, "Classification of Power Quality Events Using Optimal Time-Frequency Representations -- Part 1: Theory," IEEE Transactions on Power Delivery, 2003.

[4] M. Wang, Rowe, G. I., and A. V. Mamishev, "Classification of Power Quality Events Using Optimal Time-Frequency Representations -- Part 2: Application," IEEE Transactions on Power Delivery, 2003.

The answers are on the next page.

Answers:

In text:

- Reference [2] in the second sentence is enclosed in parentheses, whereas the standard is to enclose it in square brackets. Parentheses are used for equation numbers.
- Insertion of reference (2) in the middle of the sentence needlessly interrupts the flow of thought.
- In the third sentence, space is needed between "in" and [3,4].

In the list of references:

- Spacing between the number of the reference and the first initial of the first author is inconsistent.
- The last name "Jewell-Larsen" in reference 2 appears in Arial font. This switching of fonts is a common typesetting mistake when Reference Manager is used. It typically happens when the database line entry is copied from somewhere and the original font setting is preserved. To prevent that from happening, one can strip the font settings from text, for example, by using the command **Paste Special**, **Unformatted Text**.
- References [3] and [4] lack details, such as volume number and page numbers.
- The italics in the journal titles are inconsistent between references [1] and [2] versus [3] and [4].

EXERCISES

Exercise 4.1.

(A) Create a "master database" for your group using EndNote.
(B) Populate the master database with literature citations, numbered from 1 to 50.
(C) Create a clone database.
(D) Populate the clone database with entries from 101 to 125.
(E) Synchronize the clone and the master databases.

Exercise 4.2.

(A) Create a User-defined field.
(B) Create three categories in this field.
(C) Assign groups of papers to these categories.