

LogiTerm: Your Personal Search Engine, Part I

By Naomi J. Sutcliffe de Moraes

Please note that the reviewer is not affiliated in any way with the products being discussed. This review does not represent an endorsement of the product by ATA. All opinions are solely those of the author.

My reviews published in the August, September, and October issues focused on the environment tools Déjà Vu X and across. LogiTerm is a different kind of translation tool, with the advantage that you need not have your source text in electronic format. LogiTerm is especially useful for handwritten texts (like physicians' reports) and for specialist translators in any area.

The version reviewed here is the Professional edition, which is appropriate for freelance translators working with *languages using the Latin alphabet*. It does not work with languages using other alphabets due to encoding issues. LogiTermWebPlus does work with all languages, but costs about four times more.

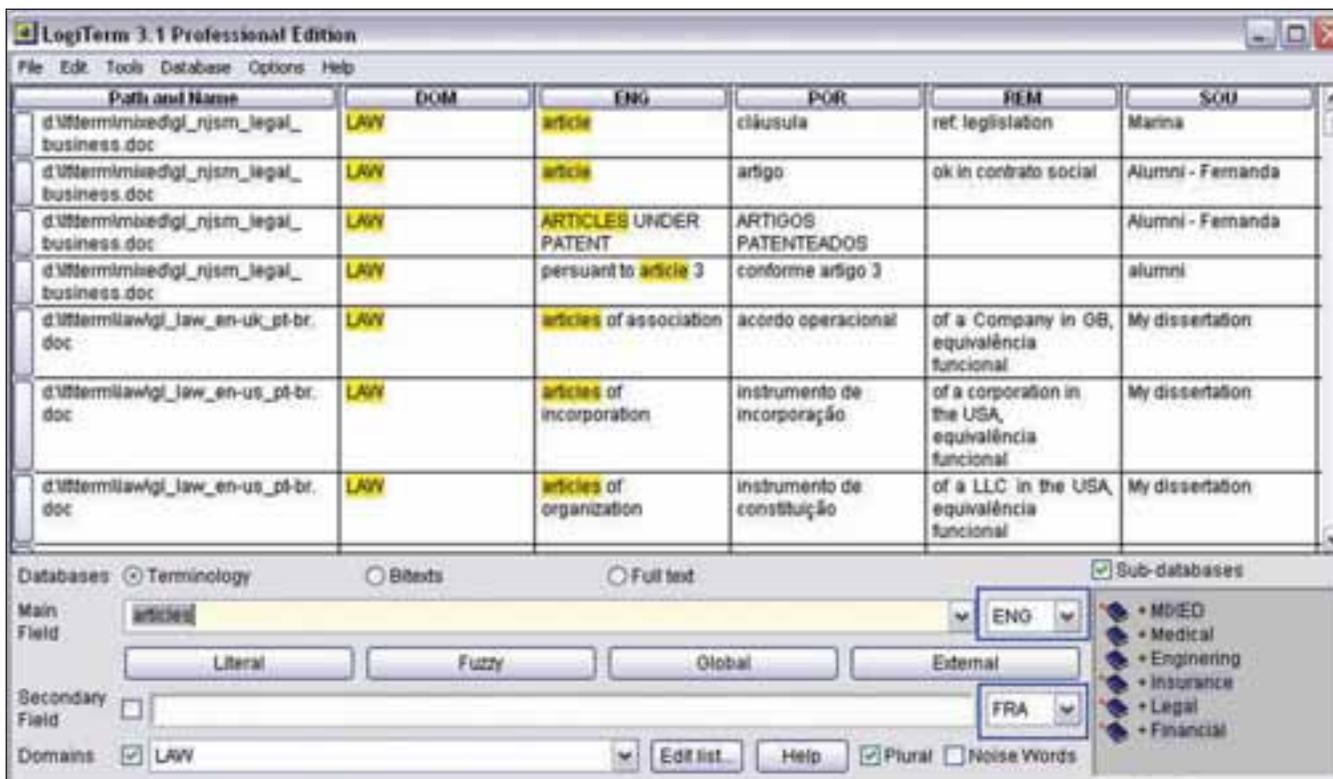
What LogiTerm Does

LogiTerm is, first and foremost, a personal search engine for translators. You define which files (on your hard drive or your network) you want to include in searches and LogiTerm indexes them. From then on, you can search these files from a single interface. This may sound just like what Google Desktop does, but it is far more useful.

LogiTerm has three databases, which are initially empty:

- 1. Terminology:** Terminology files in the native LogiTerm format and in almost any other format.
- 2. Bitexts:** Files containing originals aligned with their translations in LogiTerm format (a two-column HyperText Markup Language [HTML] file).

Figure 1: LogiTerm search interface showing fuzzy search results



3. Full Texts: Reference files in many different formats, including many Portable Document Format (PDFs).

You store your terminology, bitext, and reference files wherever you wish on your hard drive, and then tell LogiTerm where they are. After you have configured the location of the files to be searched, LogiTerm indexes them (makes a simpler internal copy of the text in the files, disregarding all formatting), which makes searching a snap.

This month's review will focus on LogiTerm's interface and use of the tool to search terminology files. Next month, we will again look at LogiTerm, this time discussing the creation of bitexts, searches in bitexts and full texts, and LogiTerm's pre-translation functions.

The Search Interface

LogiTerm provides one search interface for all three databases, as shown in Figure 1. You choose which database to search, which language, type in the search word or words, and then click on the type of search:

- **Literal:** Searches for a record matching exactly the term typed in the search field.
- **Fuzzy:** Searches for a record containing the words typed in the search field.
- **Global:** Searches for the words typed in the search field in all files indexed and in all fields in LogiTerm terminology files.

- **External:** Searches for the words typed in the search field in Termium, a French/English/Spanish online database.

This is sufficient for most searches, but you can also search for:

- Terms with (or without) specific text in a secondary field (you choose which field and which term to search on), and/or
- Terms in a specific domain (you define the domains when configuring the program), and/or
- Terms in specific sub-databases that you create.

LogiTerm has several special features that make its searches more effective:



- **Order of results:** Presents the results in order of relevance, not in some random order.
- **Plurals:** If the plural box is checked, it looks for both singular and plural forms (note that in Figure 1, I searched on the plural and it found the singular form of the word in the database). For words that do not form their plurals with an “s,” use the wildcard discussed below.
- **Articles/Prepositions:** If the noise words box is not checked, the program ignores articles and prepositions. This allows you to copy text directly from your word processor without cleaning it up. Note that

words box, so it will know to include “of” in the search.

- **Exclusion:** You can use the minus sign (-) to indicate that a specific word should not appear in the searched text. This is similar to Google, except only a terminology record or bitext segment is omitted, not an entire file.

A Place for Everything and Everything in its Place

One of the best things about LogiTerm is that it is very easy to add files/terms to the database. When you must fit your information into a specific format, you will probably end up putting it off until you have more time (which never happens). I work in four

Figure 2: My directory structure for LogiTerm files



LogiTerm is especially useful for handwritten texts (like physicians' reports) and for specialist translators in any area.

the preset noise words are French and English, but you can change them to words appropriate for your language combinations.

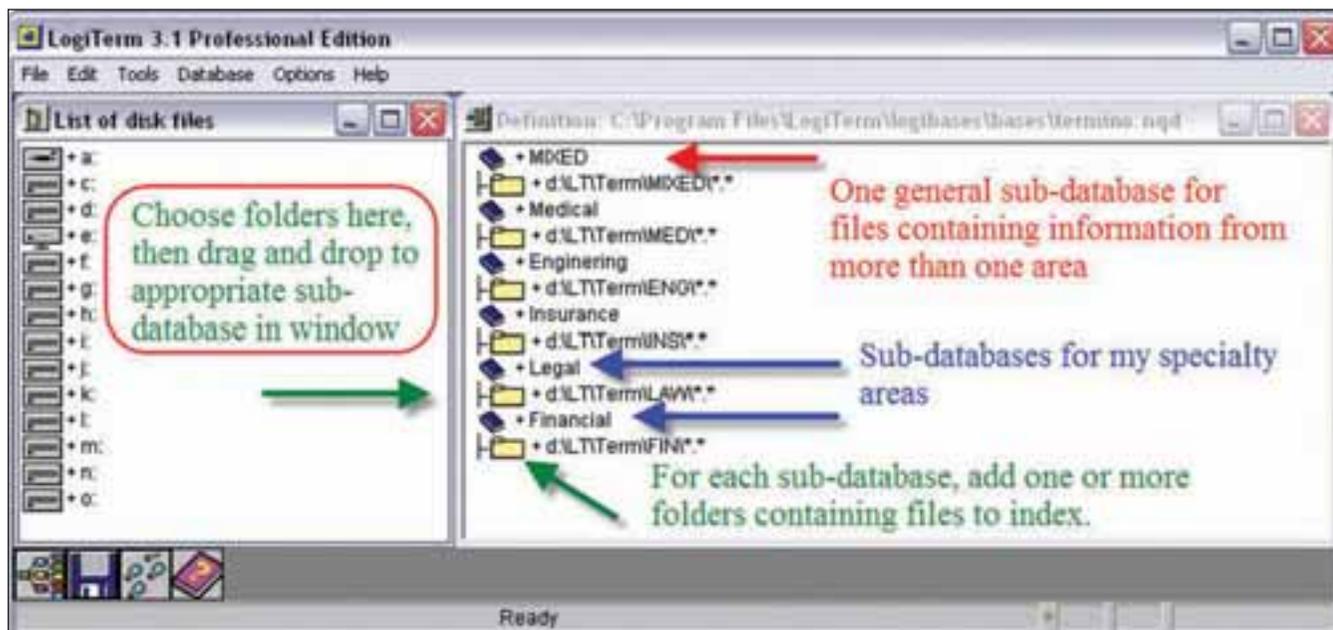
- **Wildcard:** You can use the * as a wildcard to look for all forms of a word, or for similar words. This is what I miss most when searching with Google. This is great when you cannot read the entire word when translating handwritten material.
- **Quotes:** You can use quotes to search for a string, like “articles of association” (just like in Google). Make sure you check the noise

main areas (legal, medical, engineering, and insurance), so I created four folders named LAW, MED, ENG, and INS under each of three folders for reference, bitext, and terminology files (see Figure 2). Figure 2 is just an example—you can put your files anywhere you wish and use any file structure.

Quick Inclusion: Whenever I am working and come across an interesting file or glossary, I throw it into the appropriate folder immediately (and I do not even need to open LogiTerm). If I want to be able to search the folder immediately, I update the index for that database, or I

can wait and do it later. Just this morning a client asked me to provide a quote for a file that included the articles of incorporation for a nonprofit. I collect these for reference, so I immediately saved a copy in my LT/Ref/LAW directory. Even if the client chooses a different supplier, I will have the file for my personal reference when translating similar documents. When surfing the Internet, I often run across files that would be useful either for the job I am working on or for my work in general, so now I simply download the file to the correct folder and make it part of my searchable reference material. Before I purchased LogiTerm, I often saved reference material to the client's folder, and then forgot it was there when working on similar jobs for other clients.

Figure 3: Database configuration



File Formats: Please note that LogiTerm can index (and align) MS Word, WordPerfect, HTML, MS Excel, PowerPoint, and PDF files. You do not need to convert these files before indexing. (It can process PDF files created by programs like MS Word, but not those created by scanning a document as an image.) That is why it is so easy to just drop the file into the appropriate folder without further processing. Note that you should try to give files a name that will mean something to you, since LogiTerm shows the filename and entire path in the first column of the results table (see Figure 1 on page 33).

Database Configuration: As shown in Figure 3, database configuration is very intuitive. It is very easy to set up the files and folders LogiTerm indexes, and I often create a new temporary reference database when I have a good deal of reference material for a specific project. Later, I can incorporate the specific reference material

into my area-specific database by dropping and dragging the contents of the specific folder into the regular area folders I normally use. It only takes 30

one, and the tool I use most is the terminology database. LogiTerm stores terminology records in MS Word or text files, and the interface to add new

The Professional edition is appropriate for freelance translators working with languages using the Latin alphabet.

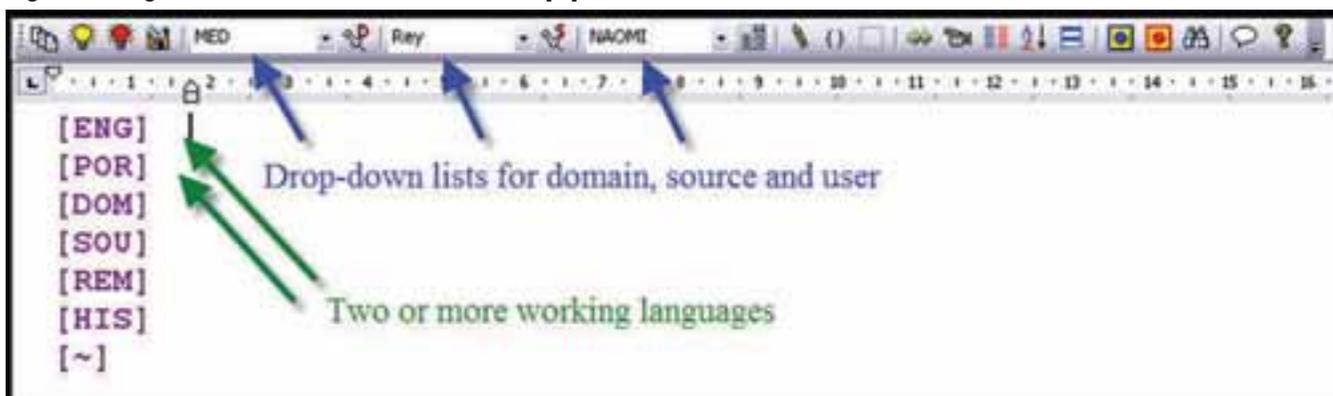
seconds to configure a new sub-database and less than a minute to update the index (though it could take longer if more than a hundred files are being indexed at once). Each database (terminology, bitexts, full texts) must be defined separately, but this also provides more flexibility.

Terminology Files

LogiTerm is really three tools in

terms is an extra toolbar in MS Word. Figure 4 on page 36 shows the LogiTerm toolbar and an empty terminology record. When I first started using LogiTerm, I was frustrated by the limited number of fields, which is one of the limitations specific to the Professional version. You can include many languages per record, but you cannot add more information fields, just domain, source, remark, ➡

Figure 4: LogiTerm toolbar in MS Word and empty record



and history (this last field is filled in automatically with the user's name and the date the record is saved). I used to keep my records in MS Excel and had at least 10 columns for each term. The extra information ended up in the remark field after conversion. Another limitation is that LogiTerm allows you to have only one template. You cannot have one English↔French database and another English↔Portuguese database. If you want three or more languages, all entries include fields for all languages (you do not need to fill them all in). Note that you can search on any word in any field, and also search on a second word in any other field (see the two drop-down lists marked with blue squares in Figure 1 on page 33).

Glossary Conversion: I have not found it difficult to transfer terminology from other formats into LogiTerm. Basically, if you can get the information into a table in MS Word with one column for one language, another column for another language, and at least one other column for everything else, you can import the information into LogiTerm's native format easily. (LogiTerm needs a tab-delimited text file, which can be created from an MS Word file, an MS Excel file, or from almost any termi-

nology tool or database.) One of the program's attractions is that it works primarily with MS Word files, with which most PC users are more familiar. The company's technical sup-

portory providing its official name and many alternative names and abbreviations for all the tests they perform. This page is 95% in Portuguese, but sometimes the English name is given

LogiTerm is, first and foremost, a personal search engine for translators.

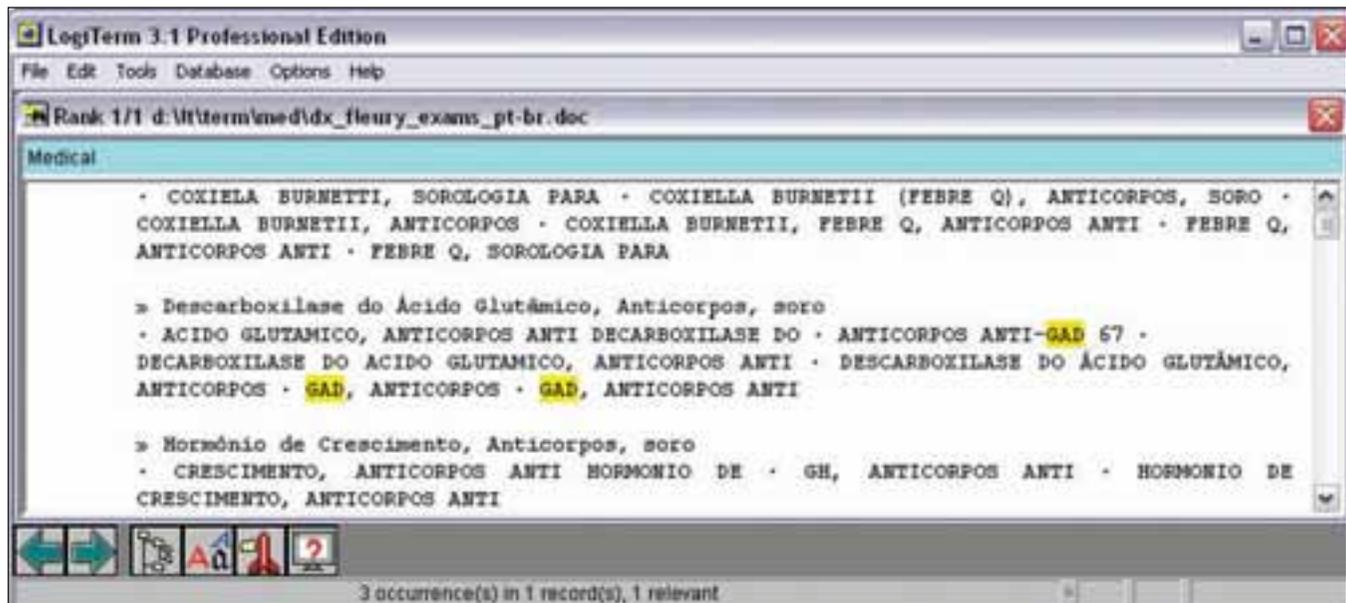
port sent me special instructions on how to convert my 10-column MS Excel data into their format shortly after I purchased the tool, and even converted one of my files for me to show me how easy it was.

Glossaries in Other Formats: Many of us collect glossaries off the Internet or from other sources, and we could spend weeks massaging them into LogiTerm's or any other program's format. This is one of LogiTerm's differentials: you can include and index any of the file types it can read without converting the data. For example, I translate medical insurance claims, and these often include laboratory reports. I found a wonderful page on the site of a Brazilian diagnostics lab-

as an alternative. I searched on the acronym GAD, which is the name of a test. The results grid shows only the name of the file—the other columns are empty. I see from the file name that the search result is probably relevant, so I click on the result and the context window in Figure 5 opens.

LogiTerm immediately shows the text in an unformatted view, with the search word highlighted in yellow. The buttons in the lower left allow you to see the text in a formatted view (which takes longer, but shows any images or formatting in the file), or can open the original file in the program in which it was created. In this example, I do not find the English equivalent, but at least I know the full name of the test. I also have a nice ref-

Figure 5: Context window



erence file containing brand names and generic names for many drugs. Terminology files do not need to be bilingual to be useful.

Other Examples: I have found LogiTerm's ability to index and find terminology in non-LogiTerm-native files invaluable. For example, a colleague sent me her rich glossary containing thousands of terms from decades of work. The only drawback was that she created the glossary in MS Word, with one column for the Portuguese term and a second column for everything else (English term, comments, source, domain, examples, etc.). It would have taken me weeks to put this in a three-column form (source, target, and everything else). Instead of converting it into LogiTerm's native format, I was able

to index it as a non-native terminology file, and now I can search it from the LogiTerm search interface. The only disadvantage is that LogiTerm has no way of knowing what is what in the file, so it cannot show me search results in the proper columns like it can with files in the native format. LogiTerm displays only the filename in the results grid and I must click on each entry to see the information in the context window, like in Figure 5.

In another case, one of my favorite engineering dictionaries came out in electronic format. Unfortunately, the CD did not contain a program with a database containing the terms, but rather a PDF file of the dictionary, but with one term per page. The only way to search for a term was to use the search function in Adobe reader, which took about five minutes to

search through all 15,000 pages. I indexed it with LogiTerm and now I can find terms instantly.

Coming Soon!

Next month, in Part II of this review, I will describe LogiTerm's automatic alignment function, how to search your bitexts, how to convert your translation memories into LogiTerm format in two easy steps, how you can collect, index, and search reference files, and how you can use LogiTerm to pre-translate source documents.

References:

LogiTerm Professional Edition v. 3.1
www.terminotix.com
 Price: 450 Canadian dollars
 (approximately US\$490)

Mark your calendar now!

ATA's 49th Annual Conference

November 5-8, 2008
 Orlando, Florida