



Millions of Terms and Translations.

Any Length. In Context.
Immediately at Your Fingertips.

MultiTrans was Designed to Address the Inherent Limitations of CAT Tools

Founder: **manager** of one of the **largest** and most advanced **multilingual publishing** and document management operations in North America;

He attempted to **improve the translation productivity** with CAT tools;

However, the available technology proved **inadequate**;

As a result he conceived **MultiTrans**, to **respond** to the inherent problems.

Experiencing the limited functionality of available Computer Assisted Translation Tools, the founder of MultiCorpora, the developers of MultiTrans, explored the possibility of using a totally different approach to address translation support and language information management.

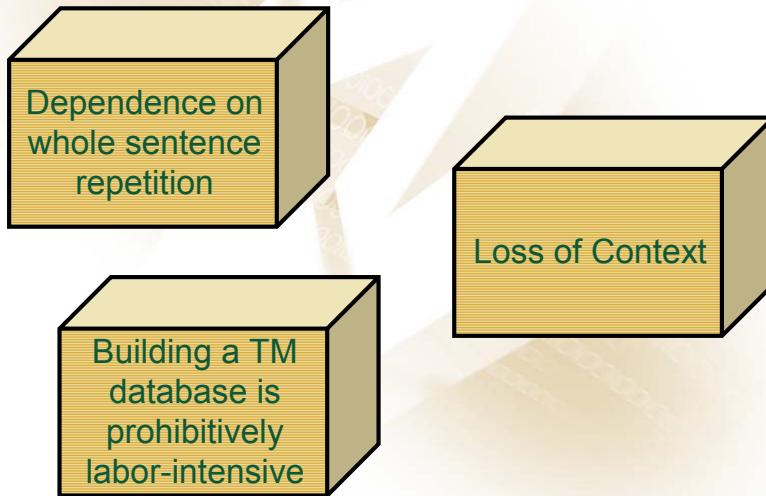
The need was to:

- Improve productivity to meet the increasing translation demand
- Improve translation speed to meet the tight deadlines
- Implement a tool that the translators would enjoy working with, and that would empower them

The current CAT tools proved **inadequate** in:

- re-using the terminological goldmine contained in a large volume of monolingual and multilingual legacy documents available;
- addressing language management needs beyond translation only. To achieve higher efficiency, a tool must be able to contribute to more aspects of the information management process;
- providing a user friendly environment for use in the day to day activities of the language professionals; and
- indexing the previous documentation take too much time, and was so tedious that it proved inefficient.

The Problems with Translation Memory



Translation Memory tools currently available have three inherent problems that significantly limit their functionality, usefulness and Return on Investment.

The Problems...



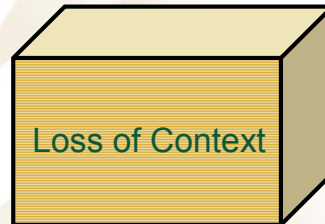
While whole sentences almost never repeat, short expressions account for the majority of the text volume in all types of content.

Not all documentation is repetitive at the sentence level, has a pre-defined life cycle and requires minor modifications every few months or years. Translators need a tool that will also allow them to experience benefits while translating non-repetitive documentation.

Studies show that most documentation, across all domains, contains repetitive expressions of 5 words or less, that must be leveraged to achieve better efficiency.

The Problems...

A lack of style and usage context results in additional time-consuming translation review and editorial work.

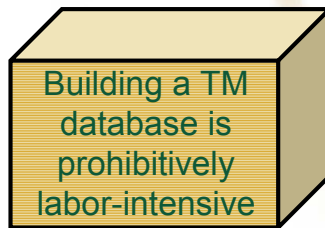


Authors, translators and editors all need contextual information to be able to make informed terminological decisions. Since conventional CAT tools work on the Translation Memory foundation which is a database of aligned sentences, retrieving contextual information (looking at the segment in its complete sentence, paragraph, or document) becomes impossible.

This is one of the main reasons why CAT tools based on a database of aligned sentences cannot address the needs of language professionals beyond the translators, such as the editors and authors.

The Problems...

It is a tedious process to perfectly align sentences with their corresponding translations, and it can take weeks to months of non income-generating time to do.



Enterprise Information Management

6

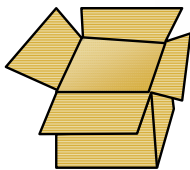
Building a database of aligned sentences is the most tedious part of implementing a conventional CAT tool. Building the Translation Memory from legacy documentation can take weeks, months and even years for larger organizations.

Furthermore, all of the sentences in your documentation need to be perfectly aligned for the Translation Memory system to work. This means that you must align an entire document perfectly, even if you may only reuse a small fraction of the sentences contained within it. Studies show that, at most, 30% of a Translation Memory actually gets re-used over time. That adds up to a lot of time wasted on alignment.

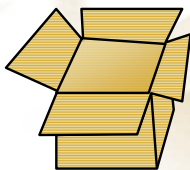
When calculating the cost of using a system, you have to consider the price of the licenses, and the implementation costs. Because building a Translation Memory of aligned sentences is so time consuming and requires non income-generating resources, this greatly decreases the potential return on investment of such a system.

The Challenge for the Next Generation of CAT Tools:

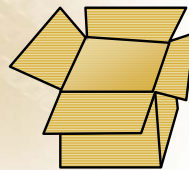
Enables the rapid creation of a vast reference pool of previous translation efforts



Provides complete usage and style context for previous translations



Effectively recycles translated expressions of any length, not just whole sentences



Multiple factors were considered in the development of the new system.

Responding to the inherent problems of Translation Memory and achieving the required milestones meant developing a more technologically advanced system than conventional Translation Memory systems, to yield better performance for the user.

Its Time for a New Approach...

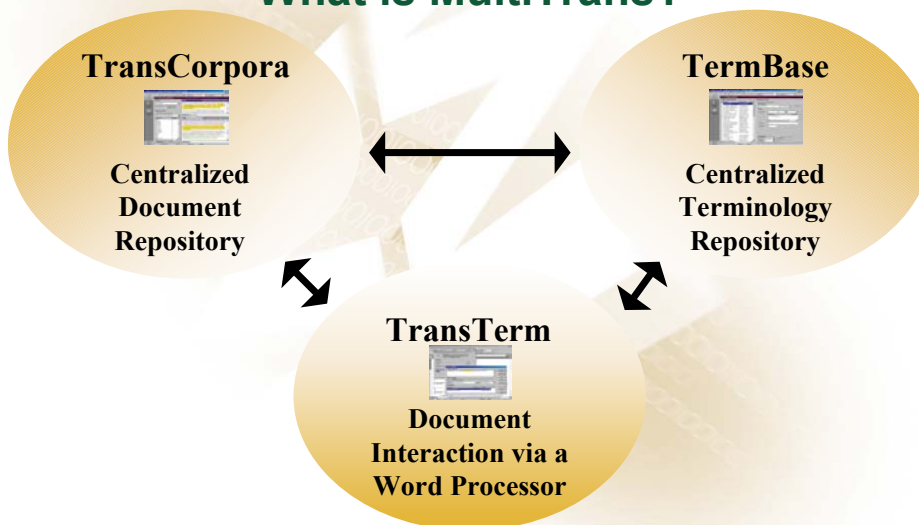
Based on full text corpus technology, MultiTrans presents a totally different solution to translation support and language management.



Millions of Terms and Translations.

Any Length. In Context. Immediately at
Your Fingertips.

What is MultiTrans?



MultiTrans integrates three separate modules:

- The TransCorpora is the centralized document repository which enables the user to quickly build and search corpora consisting of all previously translated documents.
- The TermBase is the centralized terminology repository which maintains text in its full usage context, for easy reference, management and retrieval.
- The TransTerm module incorporates all of the functionality of the TransCorpora and TermBase modules into a Word Processor.

Centralized Document Repository



- Bilingual or unilingual corpora
- Thousands of documents containing millions of previously translated terms & expressions
- Centralized approach; can be accessed on a network or intranet environment
- Allows quick and fully automated sentence alignment
- Search expressions of any length: words, phrases, sentences, paragraphs
- Terminology extraction & generation from corpus
- Easily identify & analyze words, expressions, and sentences

Instant Access to Millions of Words ...

A vast collection of legacy documents is rapidly indexed using advanced search engine techniques.

A very large searchable corpus can be built at a rate of approximately 50,000 words per minute on a low-end computer.

Any source of previously translated text can be exploited.

Translated Expressions of any Length, in their Full Context

A corpus user can perform a search of the entire corpus for an expression of any length.

The user can select and view an instance of the expression and its aligned translation in a split screen view.

At a glance the user sees the expression and aligned translation in the contexts of their complete original documents.

On-The-Fly Alignment: No Up-Front Manual Operations

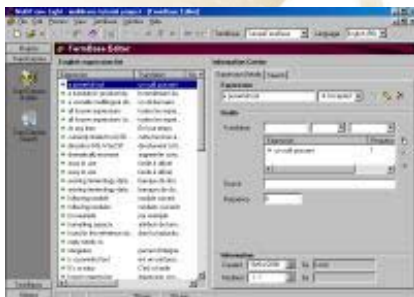
Misalignment of texts occurs frequently,
despite the sophisticated algorithms available.

A point-and-click environment in the full-text
split-screen views allows the user to easily
correct alignment on the fly.

As the user works, the corpus is continually
improving.

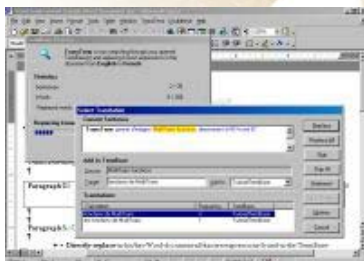
The fact that the implementation of MultiTrans requires a lot less time than the implementation of a traditional translation memory means that the return on investment will also be more appealing. Combined with the fact that MultiTrans will not only recognize full sentences, but also sub-segments, this tool brings a new level of efficiency which will also significantly reduce the pay-back period.

Centralized Terminology Repository (TermBase)



- Centralized approach; can be accessed on a network or intranet environment
- Fully customizable
- File import & export (various formats)
- Add words, expressions, sentences & even entire paragraphs
- Use many databases simultaneously!
- Flexible, concept-oriented terminology management with user-defined fields
- Scalable database architecture easily handles millions of terms

Document Interaction (TransTerm)



- Active terminology recognition and automatic propagation of translations throughout the target document
- Compare working document to all source documents in corpus (identifying corresponding segments)
- Add/replace new terms/expressions to TermBase while translating
- Interactive translation process
- Use simultaneously with other terminology banks (Termium®, etc.)
- Open architecture allows easy integration with other editing environments of your choice

The TransTerm module is the environment in which the translator will actually perform his/her work. The system interacts with the translator's preferred editing tool.

The TransTerm module will compare a document to be translated against all previously translated documents in your central document repository and identify all segments of 2 words or more that can be extracted (much like what a referencing department would do). This function can also be compared to a pro-active concordance search. It will automatically identify all concordances that can possibly be extracted from the previous documentation pro-actively.

The system will also compare a document to be translated against all opened terminology databases, and replace known terminology. This pre-translation ensures in a pro-active fashion that all approved terminology will be used by the translator, in order to ensure consistency.

This module contains a statistical analysis tool, that will compute, in a few seconds, the efficiency that you can expect to achieve by leveraging both your legacy documents and your terminology databases, greatly improving the quoting process.

Terminology

Generate Massive Multilingual Corpora!

- Terminology by example
- Terminology extraction in each language
- Rapid terminology generation (Ex.: Translation Bureau - Canada)

Terminology Management & Distribution

- Client/Server TermBase available internally or over the Internet
- Ability to massively deploy terminology over the Internet

Information Management

Built for collaborative work in a shared environment

- Create **synergy** between information-management tasks previously processed as independent entities
- **Translation** - Integral part of operational process
- **Fast searching** of multilingual information
- **Proactive** full-text multilingual search engine

Collaboration

Services

Translation

Internal Revisions

- Terminology (use and create)
- Analysis of documents yet to be translated
- Extraction of recurring expressions
- Extraction of previous translations
- Information searching

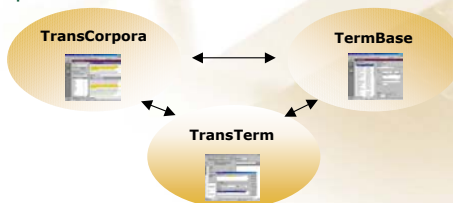
Clients

Authors/Writers

Revisors

End Users

- Terminology harmonizing
- Data exploration
- Production planning

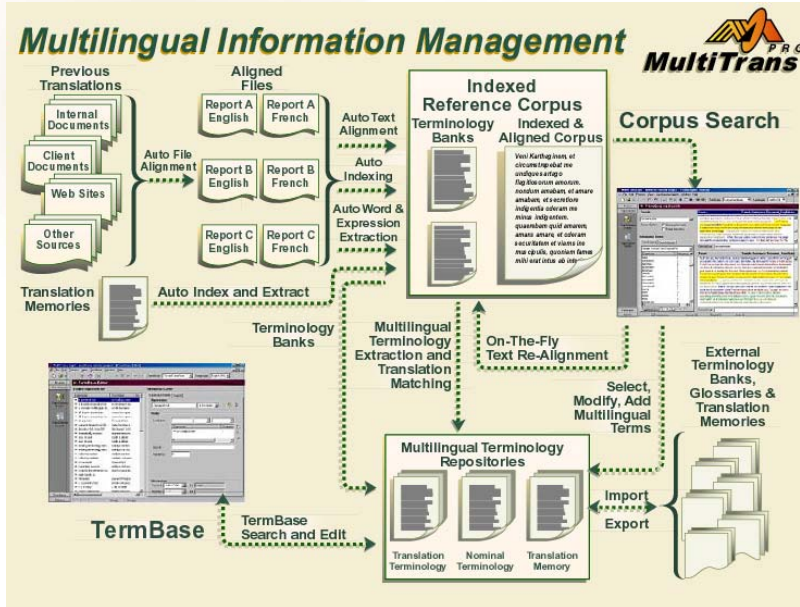


Enterprise Information Management

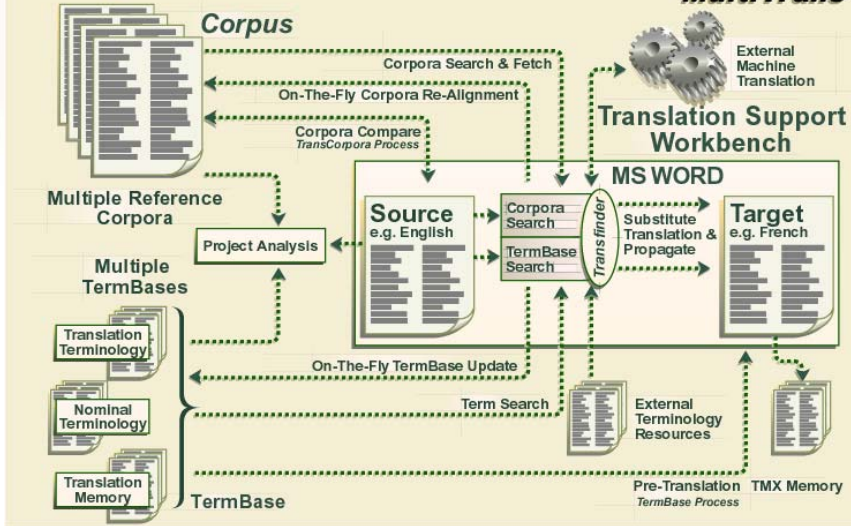
18

MultiTrans provides for an integrated collaboration between all actors in the multilingual documentation process.

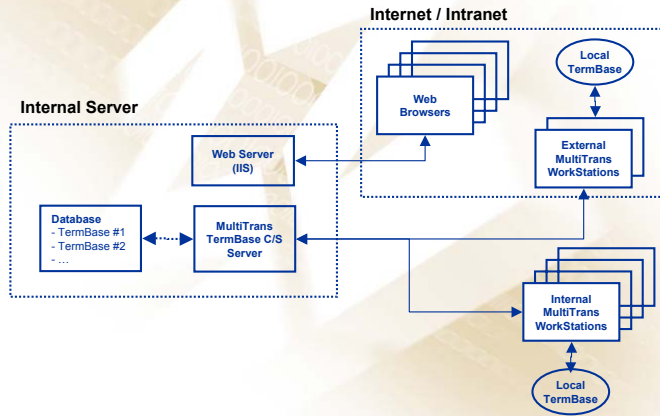
By elaborating a synergy between the authors, the translators, the terminologists, the referencing department and the editors, new levels of efficiency can be attained. This achieved synergy will improve the efficiency, productivity and quality of all departments by working from a centralized and common information base. It assures a short return on investment, while ensuring high levels of terminological consistency, both at the authoring and at the translation levels.



Integrated Translation Environment



Shared Terminology



Productivity and Quality

The user can quickly solve complex terminology and translation problems with high-quality solutions.

Up-front work is eliminated, allowing users to immediately begin exploiting the vast reservoirs of legacy translations.

The interactive nature of the technology allows the corpora and repositories to continually improve with usage.

Benefits Across the Information Management Value Chain

The vast repository of expression usage and writing style guidelines speeds up and improves the quality of monolingual content.

Information consumers have easy access to dictionaries and usage examples to solve comprehension roadblocks.

Synergies between translation and other information management processes can occur.

Conclusion

Vast reference pools of multilingual documentation
(no human-aided sentence alignment required)

Direct and rapid **creation of terminology** for specific translation projects

Scalable **corpus** (terminology by example) and **terminology bank distribution** technology

Processing at **the expression level** – go beyond conventional translation memories at the sentence or paragraph level.

Learn More

The goal of this presentation has been to briefly demonstrate the capabilities of MultiTrans. To learn more about how MultiTrans can meet your specific translation support and language management requirements, please explore some of our other available documentation:

The MultiCorpora Website: www.multicorpora.com

Breaking the Translation Memory Bottleneck: A White Paper
http://www.multicorpora.ca/papers/WhitePaper_1.pdf

The Online Product Tour
http://www.multicorpora.com/producttourpage_e.html

Contact Us

MultiCorpora R&D Inc.

CDTI de Hull
490 St. Joseph Blvd., Suite 102
Hull, Quebec, Canada
J8Y 3Y7

Tel.: +1 819.778.7070
North America Toll-free: 1 877.725.7070
Fax.:+1 819.778.0801

info@multicorpora.com
support@multicorpora.com