

DOCUMENT ALIGNMENT SOFTWARE AND WEB CRAWLER

AlignFactory Desktop and AlignFactory Robot are document alignment solutions that save significant time when feeding a computerassisted translation tool, neural machine translation engine or bilingual concordancer.

AlignFactory offers two distinct methods for pairing source and target documents: by user-defined language markers in the file names or by language detection and electronic fingerprinting algorithms that analyze file type and content. When a valid file pair is detected, AlignFactory performs the alignment and creates a bitext in the chosen format.

AlignFactory has many configuration options that let you help the software correctly identify file pairs for alignment, even if they have different file name structures. File matching criteria include:

- Language markers
- Filtering by file extension
- Exclusion strings for removing files with names containing certain character strings
- Ignore strings for ignoring certain character strings in file names
- Ignore characters after last marker
- Match files in different folders
- Allow one file name with no language marker

LogiTerm bitexts AlignFactory can create LogiTerm bitexts in XML or HTML format. If you use LogiTerm, we recommend creating XML bitexts, as they are more visually appealing and do not display segment language codes when viewed in a web browser. What's more, XML bitexts contain source document metadata and are slightly faster to index than HTML bitexts.

BENEFITS

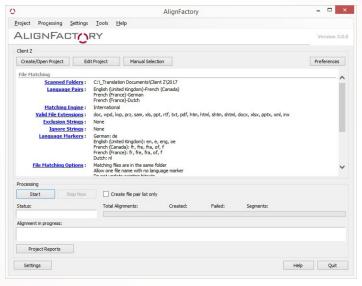
- FULLY AUTOMATED ALIGNMENT PROCESS
- DOWNLOAD AND ALIGN WEBSITES AUTOMATICALLY
- NO FILE PREP REQUIRED FOR ALIGNMENT
- XML, HTML AND TMX ALIGNMENT FORMATS
- ALIGNMENT PROJECT CREATION
- SEGMENT FILTERING
- COMPATIBLE WITH OVER 100 FILE FORMATS
- ALIGNMENT EDITOR

HTML bitexts, however, can be opened in any application with no display issues. They can also be easily indexed by any full-text search engine.

TMX files You can also create TMX files to import into any translation memory. The file creation options are as follows:

- Create one TMX file for each pair, or merge into a single file
- Insert source document name into each segment
- Automatically add attributes (project, client, domain, etc.) to segments

Alignment Editor AlignFactory has an integrated alignment editing tool that lets you make changes to LogiTerm bitexts and view and edit TMX files before importing them into your translation memory.



Main Interface

Segment filtering Over 18 filtering options that let you delete unwanted segments in your alignments automatically. Primary filters are as follows:

- Reject if both sides are the same
- Reject if segment contains no letters
- Reject duplicate segments
- Reject if too few words in segment
- Reject if one side is significantly longer
- · Reject if too many sentences in segment

Web Crawler The Web Crawler tool lets you automatically download pages or files from a website. Once the download is complete, simply create and launch an alignment project to automatically align all the downloaded content. Then, import the alignments into a computer-assisted translation tool.

The Web Crawler contains the following domain filters:

- Ignore top-level domain (TLD)
- Allow two-letter domain
- Do not truncate URL prefix

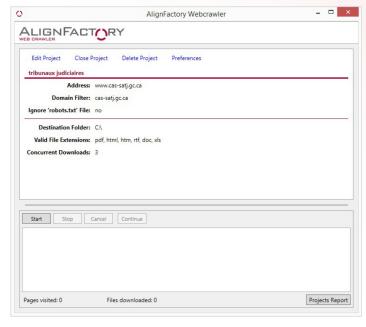
You can also filter downloaded files by extension or file name character string.

ALIGNFACT ROBOT

- Can be synchronized with LogiTerm Web and configured to scan only folders that are linked to your LogiTerm modules.
- Provides access to command-line parameters to create scripts and planned tasks, enabling the full-scale automation of the document alignment process.
- The Web Crawler module also provides access to command-line parameters to automatically update a copy of a website.
- Direct integration with LogiTerm Web, enabling you to launch alignment projects in the LogiTerm management interface.

_ 🗆 × Alignment Editor File Edit Grid Preferences The Program Review undertaken in 1994, signflicantly reduced federal spending and implied personnel reductions of a magnitude that necessitated the introduction of a variety of special departure programs to assist employees with their transition out of the public Move Cell Up Ctrl+Up Move Cell Down Ctrl+Down Le concept o Insert Row Below The concept of "payback " was developed as a performance measure to better focus accountability. comme mes la responsat tage La « période le délai néce 🖺 Copy tant This measure was defined in terms of the period of time Ctrl+C necessary to recoup the cost of the incentive programs. dépenses lié 👏 paste Ctrl+V les économia Clear Selected Cell(s) Cette mesure minanciere representant pour les minancieres representant This quantifiable measure served as a constant reminder un rappel constant de la nécessité de gérer rigoureusement les programmes de départ. to departments of the need to tightly manage the The government invested \$4.2 billion in these departure Le gouvernement a investi dans ces programmes de programs through cash settlements and increased contributions to pension funds in roughly equal proportions. départ 4,2 milliards de dollars, répartis à peu près également entre des règlements en espèces et de contributions accrues aux caisses de retraite. Jusqu'à présent, cet investissement a permis au gouvernement d'économiser plus de 8 milliards de dollars To date, the government has saved over \$8 billion as a result of salary savings resulting from this action. en salaires. The ongoing salary savings amount to \$3 billion annually, milliards de dollars par année, ce qui signfiie que le which means that the government will recoup the total XML Bitext - 574 segments Costs and savings_ENG-FRA_BT.xml

Alignment Editor



Web Crawler

TECHNICAL REQUIREMENTS

- 1 GHz processor
- 512 Mb RAM
- 140 Mb disk space
- Microsoft Windows 10 / 11 / Server 2016 / 2019 / 2022 (64 bits)
- Microsoft .NET 6



Terminotix Inc

2053 Jeanne-d'Arc Avenue, Suite 401, Montréal, Québec, Canada H1W 3Z4
T. +1514 989-9465 | sales@terminotix.com | terminotix.com



