

Canoo WMTrans Products

Canoo offers a range of smart text processing products for search engines, indexing, text mining, word stemming, spell checking, grammar checking, and machine translation applications.

WMTrans products are based on Canoo's morphological dictionary and transducer framework. The dictionary contains more than 200 000 lexeme entries and generates over 2 million words.

Deliverables

All APIs are available in Java and ANSI C++. The standard platform for the shared libraries is Linux. Other platforms are available on demand.

WMTrans Recognizer

- *Runtime program*: identifies whether or not a string is a valid German word, be it inflected or in citation form. The output is binary (t or nil).
- *Data*: data file with 100 000 or all entries.
- *Data compiler*: compiles word lists into finite state automata, thereby permitting individual extensions of the data.

WMTrans Unknown Word Recognizer

- *Runtime program*: returns whether a word is a viable non-lexicalized word formation.
- *Data*: data file with all entries and additional file with grammar rules.

WMTrans Lemmatizer

- *Runtime program*: returns the citation form of any valid German word, as used for POS taggers. Returns list of strings with Cat Features.
- *Data*: data file with 100 000 or all entries; feature file (not editable). For further information about the feature file, please refer to the documentation.
- *Data compiler*: compiles word lists into finite state automata, thereby permitting individual extensions of the data.

WMTrans Unknown Word Lemmatizer

- *Runtime program*: returns whether a word is a viable non-lexicalized word formation, propagating orthographical information (ORules) and the category.
- *Data*: data file with all entries and additional file with grammar rules.

WMTrans Inflection Analyzer

- *Runtime program*: returns the citation form and morphosyntactic classification of any valid German word, as used by language analysis programs.
- *Data*: analogous to Lemmatizer.
- *Data compiler*: analogous to Lemmatizer.

WMTrans Unknown Word Analyzer

- *Runtime program*: returns the citation form and morphosyntactic classification of a viable non-lexicalized word formation, as used by language analysis programs.
- *Data*: analogous to Lemmatizer.

WMTrans Inflection Generator

- *Runtime program*: generates all or individual inflected forms of a particular lexeme.
- *Data*: analogous to lemmatizer.
- *Data compiler*: analogous to lemmatizer.

WMTrans Inflection Analyzer/Generator

- *Runtime program*: analyzes and generates the inflected forms of a particular lexeme. This product combines the functionality of the WMTrans Analyzer and the WMTrans Generator in a single program.
- *Data*: analogous to lemmatizer.
- *Data compiler*: analogous to Lemmatizer.

WMTrans Wordformation Analyzer/Generator

- *Runtime program*: returns the base entry or entries and the word formation rule of a derived or compounded word, or all entries directly derived from a particular entry.
- *Data*: analogous to Lemmatizer.
- *Data compiler*: analogous to Lemmatizer.

Price list

Please note that the following price list is only valid for licenses to access the Canoo license server by means of the Internet.

The pricing for „offline“ licenses is determined on an individual basis. Please ask for a quote.

Licenses carry yearly license fees and include updates to new data and software releases.

The prices listed below apply to the *WMTrans Lemmatizer* without a *Data Compiler*. The Data Compiler costs an additional 10% of the total purchase price.

	No. of Accesses / second	License Cost / 20 000 entries	License Cost / 100 000 entries	License Cost / all entries (> 200 000)
Evaluation License	Unlimited	Free	n.a	n.a.
Developer License	10 / sec		600 \$US	3'000 \$US
	Unlimited (tens of thousands / sec)		1'000 \$US	5'000 \$US
Single Server Runtime Licenses (up to 5)	10 / sec		300 \$US	1'500 \$US
	Unlimited (tens of thousands / sec)		500 \$US	2'500 \$US
More than 5 Server Runtime Licenses			Ask for a quote	Ask for a quote

The prices for all other WMTrans products are based on the price for the *WMTrans Lemmatizer*.

To calculate prices for the other WMTrans products, a *functionality factor* is applied to the base price. These factors are multiplied with the corresponding price of the *WMTrans Lemmatizer*.

Product	Functionality Factor
WMTransRecognizer	0.8
WMTrans Unknown Word Recognizer	1.6
WMTrans Lemmatizer	1
WMTrans Unknown Word Lemmatizer	2
WMTrans Inflection Analyzer	1.5
WMTrans Unknown Word Analyzer	3
WMTrans Inflection Generator	1.5
WMTrans Inflection Analyzer/Generator	2
WMTrans Wordformation Analyzer/Generator	2.5

All prices are excl. VAT. The license and pricing model is subject to the respective license agreement.