

Licence Agreement

on the Use of Szeged Treebank Version 2.0

This Agreement is concluded between the
University of Szeged, Department of Informatics
H-6720 Szeged, Árpád tér 2., Hungary
<http://www.inf.u-szeged.hu/hlt>
(hereinafter referred to as **the Licensor**)

and

Name:

Position:

e-mail address:.....

Name of Workplace:.....

Address of Workplace:

Telephone:

(hereinafter referred to as **the User**)

and enters into effect as of today _____ (date: day/month/year) with the following terms and conditions:

Subject of the Licence Agreement

Szeged Treebank Version 2.0 is a data store of syntactically deep parsed texts. Szeged Treebank is one of the largest hand-processed text databases available in Hungarian, which can be used as a reference in natural language processing; furthermore, it is a good learning material for computer learning algorithms and other software applications.

The presented version of the treebank is developed on the textual basis of Szeged Corpus Version 2.0. Therefore, it incorporates texts from six different genres, with approximately 200,000 word entries from each genre, concluding 82,000 sentences in total length. Genres of Szeged Corpus 2.0 include the following:

- Fiction (*Útas és holdvilág [Journey by Moonlight]* by Antal Szerb; *Piszkos Fred, a kapitány [Dirty Fred, the Captain]* by Jenő Rejtő; and *1984* by George Orwell in full length);
- Short essays of 14 to 16 year-old students;
- Newspaper articles (excerpts from 1999 issues of three daily papers and a magazine);
- Texts related to computer science (excerpts from *Windows 2000 – haladó könyv haladó szoftverhez [Windows 2000 – advanced book on an advanced software]* by Balázs Kis and from several issues of the magazines *ComputerWorld*, *Számítástechnika [IT]*);
- Legal texts (excerpts from the 1997/144 law on economic enterprises and the 1999/76 law on author's rights);
- Short business news from the MTI (Hungarian News Agency) database.

Szeged Treebank Version 2.0 contains detailed syntactic annotation on the texts listed above. Annotation includes the marking of adjectival phrases, adverbial phrases, infinitive structures, postpositional structures, verb phrases, conjunctions and negations on top of the clause boundary annotation and noun phrase annotation already concluded on version 1.0. It further includes morpho-syntactic annotation and POS-tagging as concluded on Szeged Corpus 2.0. Texts are recorded electronically, in XML format, in compliance with TEI P4 DTD scheme.

Szeged Treebank Version 2.0 have been implemented jointly by the Department of Informatics at the University of Szeged, MorphoLogic Ltd. and the Research Institute for Linguistics at the Hungarian Academy of Sciences within the framework of the Information Technologies and Applications (IKTA) 37/2002 RTD project, supported by the Ministry of Education.

Sustaining Copyrights

Szeged Treebank Version 2.0 contains texts protected by copyright. The copyright of these texts is still in the property of their respective original owners. The Licensor declares that he has the right to make these texts available to third parties for research purposes. This does not include public reproduction, either electronically or in print, of the full text or of coherent parts of it.

The Licensor holds unlimited right of usage over Szeged Treebank Version 2.0 as a compilation and over the related annotation and documentation.

Transferring the Right of Usage: Licensing

Through the present Agreement, the Licensor grants the User a non-exclusive right of usage for the Szeged Treebank Version 2.0. The User undertakes:

- a) to use the treebank solely in a non-profit manner, for research and/or educational purposes;
- b) not to use the treebank for commercial purposes;
- c) not to modify the treebank.

The User indicates, in a clearly visible manner, the use of Szeged Treebank 2.0 in all the publications whereby he publishes results gained through the use of the treebank.

The present Licence (on the use of Szeged Treebank Version 2.0 for research and/or educational purposes) is granted by the Licensor to the User free of charge.

The commercial use of the treebank or the information gained through the use of the treebank is permitted only by the written consent of the Licensor, upon paying a licence fee. This is subject to a separate Agreement. Commercial use is not permitted under the scope of this Agreement.

The user declares that the intended use(s) of Szeged Treebank Version 2.0 is(are) the following:

- Research
- Education
- Other:

Restrictions on the Right of Usage

The User handles Szeged Treebank Version 2.0 and the related documentation confidentially, and does not disclose them to a Third Party. Exceptions are:

- a) parts of the treebank and the documentation that are expressly designed for public use;
- b) example sentences published in scientific publications and other numerical data, statistics, summary charts, etc. gained through the use of the treebank.

Apart from the above exceptions, distribution, sale, resale, leasing, disclosing to a third party, copying of the Szeged Treebank Version 2.0 is strictly **forbidden**.

The User ensures that apart from himself, only those persons have access to the files of Szeged Treebank 2.0 who are employed by the User, or work in a joint research project with the User. Every other person or institution should conclude a separate Licence Agreement with the Licensor.

Non-liability Clause

Szeged Treebank Version 2.0 is an experimental product. For this reason, the Licensor does not warrant the usability and accuracy of the treebank, and shall not be held liable for any possible damages caused by the use of the product.

The Licensor reserves the right to the continuous modification of Szeged Treebank 2.0.

Report of Errors, Privacy

In case the User finds errors and/or inconsistencies in the treebank and sends a written notification (via regular or electronic mail) to the Licensor, the Licensor guarantees that he:

- a) handles the recommendations of the User confidentially and uses it to the modification and enhancement of the treebank;
- b) discloses the suggestions of the User only to those employees – working on the development of the treebank – who are responsible for enhancing the treebank.

Termination of the Agreement

The present Licence Agreement enters into effect on the date of its signature, for an unlimited period of time. It stays effective until terminated by the Licensor or the User.

The Agreement terminates if:

- a) the User dies or its institution ceases, except if the rights and obligations of the mentioned entities are fully transferred to a successor, to whom this Agreement will apply. The User or its successor shall, in all cases, notify the Licensor of such successions. Failure to fulfil this obligation shall terminate this Agreement one month after the succession.
- b) the Licensor ceases without a successor.

In the event of the termination of this Agreement, the User shall destroy all the files of the copy of Szeged Treebank 2.0 in his possession, together with all related documentation.

Declaration

I, the undersigned User, fully understand and accept the content of the present Agreement.

.....

Date

.....

Signature