
User License for The Icelandic Gigaword Corpus, Part 1

The Árni Magnússon Institute for Icelandic Studies (hereinafter termed Operator) hereby gives the Licensee the permission to use the *The Icelandic Gigaword Corpus, Part 1* (hereinafter termed *IGCI*) in accordance with the following terms.

0 Definitions

- a) *Text* is a text in digital form, selected for use in research and the development of language technology tools.
- b) *IGCI* is a collection of selected texts that have been assembled into consistent electronic format and enhanced with morphosyntactic annotations.
- c) *The Licensee's Results* are the results of work performed by the Licensee on the texts of *IGCI*.

1 Introduction

The Árni Magnússon Institute for Icelandic Studies is the caretaker of *IGCI*.

The Árni Magnússon Institute for Icelandic Studies has reached an Agreement with Text Providers that texts from their work be a part of *IGCI*. The Text Providers have, with a separate statement, authorized the Operator to give permissions to the Licensee to use the texts as described below.

The Licensee is the user of the texts as described in this Agreement. On behalf of the Licensee there can be a group of researchers at universities or companies.

2 Terms of the Licence granted to the Licensee

- a) The operator grants a non-transferable License to the Licensee to use *IGCI* for linguistic research, for research in language technology and for the development of language technology tools. This License does not include exclusive rights to use the corpus.
- b) The access of the Licensee to *IGCI* under this License is limited to the research group of the Licensee. The research group of the Licensee consists of the Licensee's employees, contractors and other affiliates that the Licensee has authorized to carry out studies, according to article 2, paragraph *a*, with the aid of *IGCI*.
- c) *IGCI* may be installed at the place or places of work of the said research group. The place of work is defined as the computing systems that the members of a research group normally use to conduct their research activities. It can include both work and home computers, and is not restricted to particular hardware or building.
- d) Copies of *IGCI* may be made for backup purposes, or for the purposes of making data available to members of the research group, but the Licensee shall ensure that *IGCI*'s, copyright notice is reproduced on all copies of the corpus or parts thereof in accordance with article 5, paragraph *d*.
- e) Members of the said research group must not copy, publish, communicate to the public, or otherwise give to any third party access to the whole or any part of *IGCI*. *IGCI* may not be copied beyond what is identified in paragraphs *c* and *d* of this article. It is the responsibility of the

Licensee to ensure that the members of the said research group understand and abide by these restrictions and to supervise their activities with respect to *IGCI*.

- f) Neither the Licensee nor his research group may rent out or sell *IGCI*, or otherwise charge for the use of the corpus or transfer the license.
- g) There are no restrictions on the use of the Licensee's results except that the Licensee may not publish in print or electronic form or exploit commercially in any form whatsoever any extracts from *IGCI* other than those allowed under article 14 of the Copyright Act No 73/1972, and in such cases should always refer to the source.
- h) Upon publication or other disclosure or dissemination to the public that is permitted under paragraph *g* of this article, the author's moral right should be respected, according to the Copyright Act No 73/1972, and no identifiable information of third parties should be published.
- i) *IGCI* should be referenced¹ as a source in publications of research where *IGCI* has been utilized. The same applies to the release of any language technology tools that have used *IGCI*.

3 Payment

There is no charge for the use of the *IGCI* texts.

4 Period of Agreement

The License is valid indefinitely if the Licensee abides by the terms of the License. The License can be terminated as specified in article 8 should the Licensee commit any material breach of the License.

5 Requirements to the Licensee

Licensee and others in his research group must always exercise the utmost precautions to comply with the provisions of this Agreement and shall amongst other things:

- a) Maintain up-to-date written records of all copies of *IGCI* created, of all individuals in the research group of the Licensee that have access to the corpus at all times and of where these copies are stored. Records should be kept of all projects undertaken by the Licensee where *IGCI* is used. Licensee must produce such records promptly upon request from the Operator.
- b) Regularly check the status and use of *IGCI* such as updating regularly all files, according to paragraph *a*.
- c) Make use of file protection and other data security measures to prevent insofar as possible all and any unauthorized access to the files of the corpus. Access to *IGCI* shall only be allowed with the use of a password that Licensee assigns to the individuals in his research group, and this password must not be given to unauthorized persons.
- d) All necessary steps must be taken to protect copyrights, as well as intellectual property rights, owned by the Árni Magnússon Institute for Icelandic Studies, and that of the Text Providers of *IGCI*.
- e) Should the Licensee suspect his employees, members of his research group or his partners of having violated these terms or mishandled access to *IGCI*, he must immediately notify the Operator.

The Licensee shall ensure, among other things through agreements with individuals in his research group, whether employees, contractors or other partners, that they comply with the security measures

¹ Steinþór Steingrímsson, Sigrún Helgadóttir, Eiríkur Rögnvaldsson, Starkaður Barkarson and Jón Guðnason. 2018. Risamálheild: A Very Large Icelandic Text Corpus. *Proceedings of LREC 2018*, p. 4361-4366. Myazaki, Japan. 2017 version of the corpus available online at <http://malheildir.arnastofnun.is>.

specified in paragraph c, and other measures he may prescribe, and that when the research project is completed all copies of *IGCI* be deleted. An employee of the Licensee, who is so authorized, shall send the Operators a written confirmation that the Licensee has complied with this requirement.

The Licensee shall ensure that only those working on the project of the research group at any given time have access to the corpus.

6 Limitation of Liability

- a) The Operator does not guarantee that *IGCI* is suitable for a particular use or is useful under certain circumstances, even though the plans of the Licensee have previously been presented to him.
- b) The Operator is not responsible for the technical characteristics or limitations of *IGCI*.
- c) The Operator does not warrant that the operation of any software supplied with *IGCI* will be uninterrupted or free from errors.
- d) In no circumstances will *IGCI* be under any liability to the Licensee for loss of profits, goodwill or any kind of consequential losses of any nature even if such loss was foreseeable.
- e) *IGCI* shall accept no liability whatsoever in connection with any damage to the tangible property of the Licensee resulting from the negligence of the Operator, his employees, agents or subcontractors, or of other reasons.

7 Rights reserved

All intellectual property rights, including copyright, related to *IGCI*, are retained by the Text Providers and the *Árni Magnússon Institute for Icelandic Studies*.

8 Termination

- a) Operator may terminate the License by giving written notice without a special warning, if the Licensee commits any material breach of this Agreement.
- b) The License shall automatically terminate if the estate of Licensee is declared bankrupt.

9 Disputes

In the case of a dispute about this contract and if agreement is not reached between the parties, the District Court of Reykjavík has the jurisdiction.