

**Individual Agreement  
to use the  
ClueWeb12 Web Research Collections**

I, \_\_\_\_\_, am a person engaging in research and development of natural-language-processing, information-retrieval or document-understanding systems, and a member of, consultant to, or person providing service to the following organization ("Organization").

Corporation/Partnership/Legal Entity \_\_\_\_\_

Official mail address \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Telephone \_\_\_\_\_

Facsimile \_\_\_\_\_

I hereby agree to use the Information designated as the ClueWeb12 Text Research Collection(s) (the "Information"). By signing this Individual Agreement ("Agreement") with the Organization, I hereby agree to abide by the following understandings, terms and conditions. These understandings, terms and conditions apply equally to all or to part of the Information, including any updates or new versions of the Information supplied under this Agreement.

**Copyright**

1. The Information as been obtained by crawling the Internet. Due to the amount of Information it has not been practicable to obtain permission from copyright owners to provide the Information for the uses permitted under this Agreement ("Permitted Uses").
2. I understand that all the documents in the Information are documents which have been at some time made publicly available on the Internet and which have been collected using a process which respects the commonly accepted methods (such as robots.txt) for indicating that the documents should not be so collected.
3. Owners of copyright in individual documents may choose to request deletion of these documents from the Information.
4. The limitation on permitted use contained in the following section is intended to reduce the risk of any action being brought by copyright owners, but if this happens the Organization has agreed under its Organization Agreement to bear all associated liability.

**Permitted Uses**

1. The Information may only be used for research and development of natural-language-processing, Information-retrieval or document-understanding systems.

2. Summaries, analyses and interpretations of the linguistic properties of the Information may be derived and published, provided it is not possible to reconstruct the Information from these summaries.
3. Small excerpts of the Information may be displayed to others or published in a scientific or technical context, solely for the purpose of describing the research and development carried out and related issues.
4. All efforts must be made not to infringe the rights of any third party including, but limited to, the authors and publishers of any excerpts used in accordance with clause 3 above in this “Permitted Uses” section.

### **Agreement to Delete Data on Request**

I undertake to delete within thirty days of receiving notice all copies of any nominated document that is part of the Information whenever requested to do so by either:

1. Carnegie Mellon University; or
2. the owner of copyright for the particular document.

Because my Organization has entered into an Organization Agreement with Carnegie Mellon, I understand that it may receive such notices from Carnegie Mellon and/or copyright holders. In that case, I will immediately comply with any instructions from my Organization regarding the deletion of the pertinent Information.

### **Access to the Information by Individuals**

I understand that the Organization has agreed to certain obligations in respect of my access in an Agreement with Carnegie Mellon University and I agree to use the Information only for the Permitted Uses and to comply with those obligations in the Agreement as they affect me. Those obligations are that the Organization:

1. must control access to the Information by individuals and may only grant access to people working under its control, i.e., its own employees, consultants under written agreement to the Organization, or individuals providing service to the Organization under written agreement;
2. must ensure that before being given access to the Information I complete and submit this Individual Agreement form;
3. must terminate my access when I no longer require access for my work for the Organization and/or I am no longer employed by (and/or under contract with, as applicable) the Organization;
4. remains responsible for any breach of this Individual Agreement form by me;
5. shall retain the applications of all persons ever granted access to the Information and make them available upon request to any of the copyright holders and to Carnegie Mellon University;
6. shall maintain a list of people with current and recently-terminated access to the Information and make it available to Carnegie Mellon University on request; and
7. must make sure that I only display the Information to or share the Information with persons whom my Organization lists as having access to the Information.

### **No Warranty; Disclaimers**

NO WARRANTY IS GIVEN BY CARNEGIE MELLON IN RELATION TO THE INFORMATION OR THE USE(S) TO WHICH THE INFORMATION MAY BE PUT BY THE ORGANIZATION OR THE FITNESS OR SUITABILITY OF THE INFORMATION FOR ANY PARTICULAR PURPOSE OR UNDER ANY SPECIAL CONDITIONS REGARDLESS OF WHETHER ANY SUCH PARTICULAR PURPOSE OR SPECIAL CONDITIONS ARE OR HAVE BEEN MADE KNOWN TO THE UNIVERSITY PRIOR TO, OR DURING, THE PERIOD OF THIS AGREEMENT. ALL CONDITIONS AND WARRANTIES EXPRESS OR IMPLIED, WHETHER ARISING UNDER STATUTE OR UNDER COMMON LAW, INCLUDING BUT NOT LIMITED TO CONDITIONS AND WARRANTIES AS TO QUALITY, MERCHANTABILITY AND FITNESS FOR PURPOSE ARE HEREBY EXCLUDED. CARNEGIE MELLON SHALL NOT BE LIABLE TO ORGANIZATION OR ANY THIRD PARTY FOR LOSS OF PROFITS OR FOR INCIDENTAL, INDIRECT, SPECIAL OR CONSEQUENTIAL DAMAGES FOR ANY REASON WHATSOEVER ARISING OUT OF OR RELATING TO THIS AGREEMENT (INCLUDING ANY BREACH OF THIS AGREEMENT), EVEN IF CARNEGIE MELLON HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES OR HAS OR GAINS KNOWLEDGE OF THE EXISTENCE OF SUCH DAMAGES.

### **Termination**

Either Carnegie Mellon University or the Organization may terminate this Agreement at any time by notifying the other party in writing. On termination of the Agreement by either Carnegie Mellon University or the Organization or termination of my access by the Organization, I shall return to the Organization or destroy all copies of the Information.

### **Applicable Law; Disputes**

This Individual Application is governed by the laws of the state of Pennsylvania in the United States of America.

### **Miscellaneous**

If any portion of this Agreement is determined by any court or governmental agency of competent jurisdiction to violate applicable law or otherwise not to conform to requirements of law, then the rest of the Agreement will remain in effect and the parties will substitute a suitable and equitable provision for the invalid/unenforceable provision in order to carry out the original intent and purpose of the original Agreement. I may not assign any or all of my rights and/or obligations under this Agreement without the prior written consent of both Carnegie Mellon and Organization, which consent may be granted or withheld in their sole discretion. Any attempted assignment in violation of this section shall be void and of no effect. This Agreement constitutes the entire agreement and my entire understanding regarding access to and use of the Information and supersedes all previous agreements and understandings of mine relating to the subject matter of this Agreement. The Agreement may not be altered, amended, or modified except by a written instrument signed by me and by the duly authorized representatives of both Carnegie Mellon and the Organization.

I hereby execute this Agreement in favor and for the benefit of the Organization, with the understanding that Carnegie Mellon is an intended third party beneficiary of this Agreement.

**By the Individual:**

Signature \_\_\_\_\_

Date \_\_\_\_\_

Name (please print) \_\_\_\_\_

Title \_\_\_\_\_

*Completed, signed copies of this form should be retained by the Organization.*