

2012 NIST Machine Translation Open Evaluation License Agreement (MT12)

In the remainder of this document the term **User** refers to _____ of

_____ and the term **User's Research Group** refers to:

_____.

User agrees, on behalf of **User's Research Group**, to receive media (CD-ROM, DVD, hard drive, web download, etc.) containing speech and/or text data from the Linguistic Data Consortium (LDC) named below under "Corpora/Data Received" and to use the material received under this agreement (the "Data") only for purposes of the 2012 NIST Machine Translation Open Evaluation. **User** and **User's Research Group** may include limited excerpts from the Data in articles, reports and other documents describing the results of work performed in the 2012 NIST Machine Translation Open Evaluation. **User** further agrees not to otherwise publish, retransmit, disclose, display, copy, reproduce or redistribute the Data to others outside of **User's Research Group**.

User shall return the Data to LDC as follows: (1) if **User** fails to register or withdraws registration for the 2012 NIST Machine Translation Open Evaluation, **User** agrees to delete the Data and any files and software derived from it from any computer or media onto which it has been copied and to return all media to LDC **on or before April 30, 2012**; (2) if **User** fails to submit evaluation results fully and on time, **User** agrees to delete the Data and any files and software derived from it from any computer or media onto which it has been copied and to return all media to LDC **on or before April 30, 2012**; (3) if **User** fails to send a representative to the 2012 NIST Machine Translation Open Evaluation workshop, **User** agrees to delete the Data and any files and software derived from it from any computer or media onto which it has been copied and to return all media to LDC **on or before July 1, 2012**; and (4) if **User** fulfills all requirements of the 2012 NIST Machine Translation Open Evaluation, **User** agrees to delete the Data and any files and software derived from it from any computer or media onto which it has been copied and to return all media to LDC **on or before September 30, 2012**. **User** may retain the data available in LDC's general catalog by agreeing to license the data as a nonmember or by joining LDC.

All Data is provided "as is" and neither the Linguistic Data Consortium nor the University of Pennsylvania warrants the accuracy, completeness, currentness, merchantability or fitness for a particular purpose of the Data. In no event will the Linguistic Data Consortium or the University of Pennsylvania be liable for any loss or injury caused in whole or in part by its negligence or contingencies beyond its control in procuring, compiling, interpreting, editing, writing, reporting or delivering the Data, or for any errors, omissions or inaccuracies in the Data, regardless of how caused. In no event will the Linguistic Data Consortium or the University of Pennsylvania be liable to **User**, **User's Research Group**, or to any other individual or organization for any decision made or action taken in reliance upon any part of the Data or for any consequential, direct, special or similar damages, even if advised of the possibility of such damages.

User shall send a signed copy of this agreement by facsimile to the LDC, Attention: Ilya Ahtaridis, fax number (+1) 215-573-2175. Alternatively, **User** shall e-mail an electronic version of the signed agreement to the LDC at ldc@ldc.upenn.edu.

Corpora and/or Data Received:

Please indicate which corpora you would like to receive:

___	LDC2004T18	Arabic English Parallel News Part 1
___	LDC2011T11	Arabic Gigaword Fifth Edition
___	LDC2004T17	Arabic News Translation Text Part 1
___	LDC2005E46	Arabic Treebank English Translation
___	LDC2010T13	Arabic Treebank: Part 1 v 4.1
___	LDC2011T09	Arabic Treebank: Part 2 v 3.1
___	LDC2010T08	Arabic Treebank: Part 3 v3.2
___	LDC2010L01	LDC Standard Arabic Morphological Analyzer (SAMA) Version 3.1
___	LDC2005T34	Chinese <-> English Name Entity Lists (v1.0)
___	LDC2005T10	Chinese English News Magazine Parallel Text
___	LDC2009T27	Chinese Gigaword Fourth Edition
___	LDC2005T06	Chinese News Translation Text Part 1
___	LDC2005T23	Chinese Proposition Bank 1.0
___	LDC2010T07	Chinese Treebank 7.0
___	LDC2003E07	Chinese Treebank English Parallel Corpus
___	LDC2002L27	Chinese-English Translation Lexicon Version 3.0
___	LDC2011T07	English Gigaword Fifth Edition
___	LDC2004E72	eTIRR Arabic English News Text
___	LDC2003E14	FBIS Multilanguage Texts
___	LDC2007E06	GALE Phase 2 Release 1 - Translations
___	LDC2007E46	GALE Phase 2 Release 2 - Translations
___	LDC2007E87	GALE Phase 2 Release 3 - Translations
___	LDC2007E101	GALE Phase 3 Release 1 - Translations
___	LDC2008E40	GALE Phase 3 Release 2 - Translations
___	LDC2008E56	GALE Phase 4 Release 1 - Translations V2.0
___	LDC2009E16	GALE Phase 4 Release 2 - Translations
___	LDC2009E95	GALE Phase 4 Release 3 - Translations V1.2
___	LDC2004T08	Hong Kong Parallel Text
___	LDC2006T14	Korean Broadcast News Transcripts
___	LDC2010T19	Korean Newswire Second Edition
___	LDC2003T08	Korean Telephone Conversations Transcripts
___	LDC2006T04	Multiple Translation Chinese (MTC) Part 4
___	LDC2003T18	Multiple-Translation Arabic (MTA) Part 1
___	LDC2005T05	Multiple-Translation Arabic (MTA) Part 2
___	LDC2003T17	Multiple-Translation Chinese (MTC) Part 2
___	LDC2004T07	Multiple-Translation Chinese (MTC) Part 3
___	LDC2002T01	Multiple-Translation Chinese Corpus
___	LDC2006E44	TIDES MT 2004 Arabic evaluation data
___	LDC2006E43	TIDES MT 2004 Chinese evaluation data
___	LDC2006E39	TIDES MT 2005 Arabic evaluation data
___	LDC2006E38	TIDES MT 2005 Chinese evaluation data
___	LDC2004E13	UN Arabic English Parallel Text
___	LDC2004E12	UN Chinese English Parallel Text
___	LDC2002E18	Xinhua Chinese English Parallel News Text Version 1 beta
___	LDC2000T50	Hong Kong Hansards Parallel Text
___	LDC2000T47	Hong Kong Laws Parallel Text
___	LDC2000T46	Hong Kong News Parallel Text

___	LDC2007T08	ISI Arabic-English Automatically Extracted Parallel Text
___	LDC2007T09	ISI Chinese-English Automatically Extracted Parallel Text
___	LDC2011T03	OntoNotes Release 4.0
___	LDC2007T03	Tagged Chinese Gigaword
___	LDC2003T07	Arabic Treebank: Part 1 - 10K-word English Translation
___	LDC2007T02	English Chinese Translation Treebank v 1.0
___	LDC2006T10	English-Arabic Treebank v 1.0
___	LDC2006T13	Web 1T 5-gram Version 1
___	LDC2005E83	GALE Y1 Q1 Release - Translations
___	LDC2006E85	GALE Y1 Q3 Release - Translations
___	LDC2006E92	GALE Y1 Q4 Release - Translations
___	LDC2006E24	GALE Y1 - Interim Release: Translations
___	LDC2006E34	GALE Y1 Q2 Release - Translations V2.0
___	LDC2010T17	NIST 2006 Open Machine Translation (OpenMT) Evaluation
___	LDC99T42	Treebank-3
___	LDC2006E26	GALE Financial News
___	LDC2002T31	AQUAINT
___	LDC2005T28	HARD
___	LDC2005T35	ANC
___	LDC93T1	ACL/DCI
___	LDC93T3A	TIPSTER
___	LDC95T21	NAN
___	LDC95T6	CSR-III
___	LDC98T30	NAN Supplement
___	LDC2005E85	GALE Y1 Q1 Release - English Translation Treebank
___	LDC2006E25	GALE Y1 - Arabic English Parallel News Text
___	LDC2006E36	GALE Y1 Q2 Release - English Translation Treebank
___	LDC2006E82	GALE Y1 Q3 Release - English Translation Treebank
___	LDC2006E95	GALE Y1 Q4 Release - English Translation Treebank
___	LDC2006E86	GALE Y1 Q3 Release - Word Alignment
___	LDC2006E93	GALE Y1 Q4 Release - Word Alignment
___	LDC2010T23	NIST 2009 Open Machine Translation (OpenMT) Evaluation
___		OpenMT12 Arabic-to-English Evaluation Data
___		OpenMT12 Chinese-to-English Evaluation Data
___		OpenMT12 Dari-to-English Evaluation Data
___		OpenMT12 Farsi-to-English Evaluation Data
___		OpenMT12 Korean-to-English Evaluation Data

Organization: _____

Name: _____

Mailing address: _____

Phone: _____

Email: _____

Signature: _____

Date: _____

For LDC:

Christopher Cieri
Executive Director