

2015 NIST Machine Translation Open Evaluation License Agreement (MT15)

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 2015 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 2015 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 2015 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 January 15, 2015**; (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 February 28, 2015**; (3) if **User** fails to send a representative to the 2015 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 June 30, 2015**; and (4) if **User** fulfills all requirements of the 2015 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, 2015**. **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
___	LDC2012T07	Arabic Treebank – Broadcast News v1.0
___	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
___	LDC2011T11	Chinese Gigaword Fifth Edition
___	LDC2005T06	Chinese News Translation Text Part 1
___	LDC2005T23	Chinese Proposition Bank 1.0
___	LDC2008T07	Chinese Proposition Bank 2.0
___	LDC2013T13	Chinese Proposition Bank 3.0
___	LDC2013T21	Chinese Treebank 8.0
___	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
___	LDC2004T08	Hong Kong Parallel Text
___	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
___	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
___	LDC2009T14	Tagged Chinese Gigaword Version 2
___	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
___	LDC99T42	Treebank-3
___	LDC2002T31	AQUAINT
___	LDC2005T28	HARD
___	LDC2005T35	ANC

___	LDC93T1	ACL/DCI
___	LDC93T3A	TIPSTER
___	LDC95T21	NAN
___	LDC95T6	CSR-III
___	LDC98T30	NAN Supplement
___	LDC2007T24	GALE Phase 1 Arabic Broadcast News Parallel Text - Part 1
___	LDC2007T23	GALE Phase 1 Chinese Broadcast News Parallel Text - Part 1
___	LDC2008T02	GALE Phase 1 Arabic Blog Parallel Text
___	LDC2008T06	GALE Phase 1 Chinese Blog Parallel Text
___	LDC2008T08	GALE Phase 1 Chinese Broadcast News Parallel Text - Part 2
___	LDC2008T09	GALE Phase 1 Arabic Broadcast News Parallel Text - Part 2
___	LDC2008T18	GALE Phase 1 Chinese Broadcast News Parallel Text - Part 3
___	LDC2009T09	GALE Phase 1 Arabic Newsgroup Parallel Text - Part 2
___	LDC2009T02	GALE Phase 1 Chinese Broadcast Conversation Parallel Text - Part 1
___	LDC2009T03	GALE Phase 1 Arabic Newsgroup Parallel Text - Part 1
___	LDC2009T06	GALE Phase 1 Chinese Broadcast Conversation Parallel Text - Part 2
___	LDC2009T15	GALE Phase 1 Chinese Newsgroup Parallel Text - Part 1
___	LDC2010T03	GALE Phase 1 Chinese Newsgroup Parallel Text – Part
___	LDC2012T06	GALE Phase 2 Arabic Broadcast Conversation Parallel Text Part 1
___	LDC2012T14	GALE Phase 2 Arabic Broadcast Conversation Parallel Text Part 2
___	LDC2012T17	GALE Phase 2 Arabic Newswire Parallel Text
___	LDC2012T18	GALE Phase 2 Arabic Broadcast News Parallel Text
___	LDC2013T01	GALE Phase 2 Arabic Web Parallel Text
___	LDC2013T11	GALE Phase 2 Chinese Broadcast Conversation Parallel Text Part 1
___	LDC2013T16	GALE Phase 2 Chinese Broadcast Conversation Parallel Text Part 2
___	LDC2014T04	GALE Phase 2 Chinese Broadcast News Parallel Text Part 1
___	LDC2014T11	GALE Phase 2 Chinese Broadcast News Parallel Text Part 2
___	LDC2013T10	GALE Arabic-English Parallel Aligned Treebank – Newswire
___	LDC2012T16	GALE Chinese-English Word Alignment and Tagging Training Part 1
___	-- Newswire and Web	
___	LDC2012T20	GALE Chinese-English Word Alignment and Tagging Training Part 2
___	-- Newswire	
___	LDC2012T24	GALE Chinese-English Word Alignment and Tagging Training Part 3
___	-- Web	
___	LDC2013T05	GALE Chinese-English Word Alignment and Tagging Training Part 4
___	-- Web	
___	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
___	LDC2010T10	NIST 2002 Open Machine Translation (OpenMT) Evaluation
___	LDC2010T11	NIST 2003 Open Machine Translation (OpenMT) Evaluation
___	LDC2010T12	NIST 2004 Open Machine Translation (OpenMT) Evaluation
___	LDC2010T14	NIST 2005 Open Machine Translation (OpenMT) Evaluation
___	LDC2010T17	NIST 2006 Open Machine Translation (OpenMT) Evaluation
___	LDC2010T21	NIST 2008 Open Machine Translation (OpenMT) Evaluation
___	LDC2010T23	NIST 2009 Open Machine Translation (OpenMT) Evaluation
___	LDC2013T03	NIST 2012 Open Machine Translation (OpenMT) Evaluation
___	LDC2013T07	NIST 2008-2012 Open Machine Translation (OpenMT) Progress Test
___	Sets	

___	LDC96S49	CALLFRIEND Egyptian Arabic
___	LDC96S55	CALLFRIEND Mandarin Chinese-Mainland Dialect
___	LDC96S56	CALLFRIEND Mandarin Chinese-Taiwan Dialect
___	LDC97S45	CALLHOME Egyptian Arabic Speech
___	LDC2002S37	CALLHOME Egyptian Arabic Speech Supplement
___	LDC96S34	CALLHOME Mandarin Chinese Speech
___	LDC97T19	CALLHOME Egyptian Arabic Transcripts
___	LDC2002T38	CALLHOME Egyptian Arabic Transcripts Supplement
___	LDC96T16	CALLHOME Mandarin Chinese Transcripts
___	LDC2013E81	BOLT Phase 2 Translation Training Data R1
___	LDC2013E85	BOLT Phase 2 Translation Training Data R2
___	LDC2013E118	BOLT Phase 2 Translation Training Data R3
___	LDC2013E125	BOLT Phase 2 Translation Training Data R4
___	LDC2013E132	BOLT Phase 2 Translation Training Data R5
___	LDC2014E18	BOLT Phase 2 Translation Training Data R6
___	LDC2013E80	BOLT Phase 2 Translation DevTest Data R1
___	LDC2013E83	BOLT Phase 2 Translation DevTest Data R2
___	LDC2014E09	BOLT Phase 2 Translation DevTest Data R3
___	LDC2013E119	BOLT Phase 2 SMS and Chat DevTest Gold Standard Translation
___	LDC2014E25	BOLT Phase 2 SMS and Chat DevTest Gold Standard Translation R2
___	LDC2013E94	BOLT Phase 2 Arabic CTS Translation Data R1
___	LDC2014E08	BOLT Phase 3 Translation Training Data R1
___	LDC2014E50	BOLT Phase 3 Translation Training Data R2
___	LDC2014E69	BOLT Phase 3 Translation DevTest Data R1
___	LDC2014E86	BOLT Phase 3 Translation DevTest Data R2
___	LDC2014E99	BOLT Phase 3 Translation Training Data R3
___	LDC2014E110	BOLT Phase 3 Translation Training Data R4
___	LDC2014E111	BOLT Phase 3 CTS DevTest Gold Standard Translation
___	NIST 2015 Open Machine Translation (OpenMT) Evaluation Set (to be sent by NIST)	

Organization: _____

Name: _____

Mailing address: _____

Phone: _____

Email: _____

Signature: _____

Date:

For LDC:

Christopher Cieri
Executive Director