

## IWSLT 2008 Corpus - End-User Agreement

### INSTRUCTIONS for COMPILATION

Please compile and submit this End-User Agreement in two original signed copies (this page is NOT part of the agreement).

Please complete the reference (Reference: \_\_\_\_\_) at the top of page 1 of the contract with the name of your institution.

Each page of the agreement should also be initialed (the person signing the agreement should put his/her initials on each and all pages).

The END-USER named must be a legal institution, or a department or section of a named legal institution, not an individual nor a project - the person signing this Agreement should be the person authorized by the institution for such signatures (e.g. Department or Administrative Head or similar).

In Exhibit B, you must insert the address of the location where the data will be used.

*All forms must be sent by mail to:*

- *Michael Paul – IWSLT 2008 Evaluation Campaign Chair.*

National Institute of Information and Communications Technology

MASTAR Project

Language Translation Group

2-2-2 Hikaridai, Keihanna Science City

Kyoto 619-0288, Japan

# International Workshop on Spoken Language Translation

## IWSLT 2008 Corpus

### END-USER AGREEMENT

(Reference: \_\_\_\_\_)

**This agreement is made by and between:**

.....  
(hereinafter called END-USER), having its principal place of business at:

.....  
AND

**Advanced Telecommunications Research Institute International**, acting on behalf of the C-Star consortium, and hereinafter called ATR, having its principal place of business at: *2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan*, a cooperation organized and existing under the laws of Japan.

**whereby it is agreed as follows:**

1. The **IWSLT 2008 Corpus**, object of this Agreement, for which ATR, C-Star Consortium members obtained distribution rights from the rightful holders (hereinafter called PROVIDERS) for the use within the 2008 IWSLT evaluation campaign, is described in Exhibit A.
2. The site of use of the **IWSLT 2008 Corpus** is mentioned in Exhibit B.
3. Within this Agreement ATR grants END-USER, engaged in *bona fide* language engineering research, the non-exclusive right to use the **IWSLT 2008 Corpus**, exclusively for the purposes of linguistic research.
4. END-USER is not permitted to reproduce the **IWSLT 2008 Corpus** for commercial or distribution purposes and to commercialize (or distribute for free) in any form or by any means the **IWSLT 2008 Corpus** or any derivative product or services based on all or a substantial part of it. 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. 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 and related issues. Any such use shall not infringe on the rights of any third party including, but limited to, the authors and publishers of the excerpts.
5. ATR grants END-USER the right to reproduce the **IWSLT 2008 Corpus** temporarily or permanently, to translate, adapt, arrange and modify by any means the **IWSLT 2008 Corpus**, if one or all of these acts are necessary to access and make the agreed use of the contents of the **IWSLT 2008 Corpus**. END-USER is permitted to make a copy of the **IWSLT 2008 Corpus** for purposes of archiving only.
6. Without prejudice to the other provisions, the rights referred to herein shall be non transferable to any other entity. The **IWSLT 2008 Corpus** shall not be transferred to or accessed from any other site.
7. END-USER acquires no ownership, rights or title in all or any parts of the **IWSLT 2008 Corpus**.
8. ATR and PROVIDERS accept no responsibility for the accuracy or completeness of the data or for the consequences of their use. ATR and PROVIDERS give no warranty for merchantability and/or fitness for a particular purpose of the **IWSLT 2008 Corpus**.

9. END-USER shall give appropriate references to IWSLT 2008 in scholarly literature when the **IWSLT 2008 Corpus** are mentioned.
10. END-USER shall not use the name of ATR nor the names of PROVIDERS in any publication in any manner that would imply an endorsement of END-USER or any product or service offered by END-USER.
11. END-USER has no right or authority to incur, assume or create, in writing or otherwise, any warranty, liability or other obligation of any kind, express or implied, in the name of or on behalf of ATR, it being intended that each party shall remain an independent contractor responsible for its own actions.
12. Neither party shall be responsible for damages caused by the other party's non-fulfillment of provisions of this AGREEMENT.
13. Both parties exclude all liability of whatsoever nature for direct, consequential or indirect loss or damage suffered by the other.
14. ATR agrees to provide **IWSLT 2008 Corpus** to END-USER free of charge.
15. END-USER will maintain and post a list of people with current and recently-terminated access to the **IWSLT 2008 Corpus**. END-USER shall maintain such appropriate records and ATR shall have the right to audit such records upon reasonable prior notice using an independent auditor of its choice (the "Auditor"). The Auditor shall be bound to keep confidential the details of the business affairs of END-USER.
16. END-USER agrees not to disclose nor communicate to anyone any information concerning or contained in technical or commercial documents relating to the IWSLT 2008 Corpus placed at END-USER's disposal by ATR, except to its officers, directors or employees on a "need-to-know" basis.
17. In order to preserve the desired pre- competitive nature of the IWSLT conferences, the C-Star consortium has developed a set of guidelines constraining the dissemination and publication of IWSLT evaluation results. These guidelines are meant to preclude the publication of incomplete or inaccurate information that could damage the reputation of the conference or its participants and could discourage participation in future conferences. The guidelines (listed in Exhibit C) shall be implemented by END-USER.
18. This AGREEMENT is subject to, construed and interpreted in accordance with the Law of Japan. Should it not be possible to settle amicably differences of interpretation out of this Agreement, then the case shall be brought before the regular Japanese courts of law for a decision. This AGREEMENT may be terminated by ATR if END-USER breaches any material term or provision of this AGREEMENT, including Article 18. In the event that one of these articles is considered by Japanese law or a Japanese court decision as null, this invalid or illegal article shall not affect validity, legality and enforceability of the remaining provisions.
19. The term of validity of this AGREEMENT shall be from the date this AGREEMENT has been concluded until October 21, 2009. The agreement shall be renewable at the end of the current term for a successive one year unless either the END-USER or ATR gives written notice of its intension not to renew the agreement before expiration of the current term. The **IWSLT 2008 Corpus** may only be used during the term of validity of this AGREEMENT.

The entire AGREEMENT is composed of the 19 articles herein together with Exhibits A, B, and C thereafter.

In witness hereof, intending to be bound, the parties hereto have executed this AGREEMENT by their duly authorized officers.

AUTHORISED BINDING SIGNATURES:

\_\_\_\_\_  
On behalf of  
Name:  
Title:  
Date:

\_\_\_\_\_  
On behalf of ATR  
Name: Michael Paul  
Title: IWSLT 2008 Evaluation Campaign Chair  
Date:

## EXHIBITS

### *Exhibit A: IWSLT 2008 Corpus*

IWSLT 2008 Corpus refers to the resources listed below:

(Supplied Corpora provided to participants of IWSLT 2008)

- **Arabic corpus:** InterACT, Carnegie Mellon University, copyright 2005-2008
- **Chinese corpus:** Chinese Academy of Science (CAS), copyright 2004-2008  
(evaluation data sets) ATR, copyright 2006-2008
- **English corpus:** ATR, copyright 2004-2008
- **Spanish corpus:** ATR, copyright 2008.

(MT System Outputs and Subjective Evaluation Results of all systems participating in IWSLT08)

- **CHALLENGE Task:** Chinese-to-English, English-to-Chinese
- **BTEC Task:** Arabic-to-English, Chinese-to-English, Chinese-to-Spanish
- **PIVOT Task:** Chinese-to-English-to-Spanish

### *Exhibit B: SITE OF USE (insert here below the address of the location where the data will be used)*

.....

.....

.....

.....

.....

.....

### *EXHIBIT C: Guidelines constraining the dissemination and publication of 2008 IWSLT evaluation results are:*

- a) **SCIENTIFIC OR TECHNICAL PUBLICATIONS:** Scientific or technical publications, including newsletters from universities or research laboratories, should adhere to community standards for fairness and objectivity and should accurately and clearly state the limitations of the testing conditions and other factors which might influence scores. The experimental nature of the tasks, data and evaluation procedures should also be stated. The full conference proceedings should always be referenced.
- b) **ADVERTISEMENTS:** No advertisements using the 2008 IWSLT evaluation results can be placed in magazines, journals, newspapers, or other publications.
- c) **PRESS RELEASES:** Press releases about the 2008 IWSLT results to organizations with national/international coverage are also prohibited.
- d) **MARKETING LITERATURE, LOCAL NEWSLETTERS:** Although it is recognized that extensive evaluation discussions are not appropriate in this type of literature, it is expected that any claims made on the basis of evaluation results are accurate, that the evaluation measures used to substantiate these claims are stated, and that a reference is made to the full conference proceedings. Where promotional material is subject to prepublication revision by the media, the author should make every effort to see that the revision does not cause a violation of the guidelines.
- e) **CROSS-SYSTEM COMPARISONS:** Cross-system comparisons may not be made with other named teams for individual tests, and may only be made when they are supported by accepted methods of statistical significance testing. Comparisons must be accompanied by the results of those tests and should reference the publication of those tests in the conference proceedings or related literature. Informal, qualitative comparisons with recognized baselines or benchmarks, and with general levels or trends in performance, must be clearly stated to be such and thus open to statistical reassessment.